

DR GO; GO:0015829; F-lyase activity; IEA.
 DR GO; GO:000152; P-metabolism; IEA.
 DR InterPro; IPR005630; Terpene synth_C.
 DR InterPro; IPR008949; Terpenoid synth.
 DR InterPro; IPR008930; Terp_cyc toroid.
 DR InterPro; IPR001906; Terp synth-like.
 DR Pfam; PF01397; Terpene synth; 1.
 DR Pfam; PF03936; Terpene synth_C; 1.
 KW Transit peptide.
 FT TRANSIT 1 99 POTENTIAL.
 FT CHAIN 100 599 4S-LIMONENE SYNTHASE.
 SQ SEQUENCE 599 AA; 69843 MW; 8BF53D32329CAF65 CRC64;

Alignment Scores:
 Pred. No.: 6,648-218 Length: 599
 Score: 2876.00 Matches: 542
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.90% Indels: 0
 DB: 10 Gaps: 0

US-09-938-956-6 (1-1632) x Q40322 (1-599)

QY	4	AGACGATCGGAAACTCAACACCTCTCGTTGGGATGTCACCTTCATCCAAATCGCTTC	63
Db			
Db	58	ArgArgSerGlyAsnTyrAsnProSerArgTrpAspValAsnPheIleGlnSerLeuLeu	77
QY	64	AGTGACTATAGGAGGACAAACACGATGATGAGGCTTCTGAGTGGTCACTTTGGTGAAG	123
Db			
Db	78	SerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuValLys	97
QY	124	ATGGAACTCGGAAAGAAACGGATCAAAATTCGACAACTTCGATGATCGATGACTTCGAC	183
Db			
Db	98	MetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspAspLeuGln	117
QY	184	AGGATGGGCTGTCGATCATTCCAAATGATGTTCAAGAAATCTTGCTCTATATAT	243
Db			
Db	118	ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluLeuLeuSerIleTyr	137
QY	244	CTCGACCATCACTATTACAAAGACCCCTTTTCCAAAGAAAGAGGATCTCTACTCCACA	303
Db			
Db	138	LeuAspHisIleTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr	157
QY	304	TCTCTTGCACTTAGCTCTCAGAGACATGTTTCACTCCGACACAGAGGTATTCGAT	363
Db			
Db	158	SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp	177
QY	364	AGTTTCAAGAACGAGGAGGTGAGTTCAAAGAAAGCCTTAGCGACGACACACGAGGATTG	423
Db			
Db	178	SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu	197
QY	424	TTGCAACTGTATGAGCTTCTCTGTTGACGAGCGAAGCCACCGCTCGAGTCAGCG	483
Db			
Db	198	LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrLeuGluSerAla	217
QY	484	AGGGAATTCGCCACAAATTTTGGAGGAAAGTGAACGAGGTGTTGATGCGAC	543
Db			
Db	218	ArgGluPheAlaThrLysPheLeuGluLysValAsnGluGlyValAspGlyAsp	237
QY	544	CTTTTAAACAGAAATCGCATATCTTGGACATCCCTCTTCATGGAGGATTAAGGCCA	603
Db			
Db	238	LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro	257
QY	604	AATGCACCTGTGTGATCGAATGGTATGGAAGAGCGCCGACATGAATCCAGTGTGTG	663
Db			
Db	258	AsnAlaProValTrpIleGluTrpArgLysArgProAspMetAsnProValValLeu	277
QY	664	GAGCTTGCATACCTCGACTTAAATATGTTCAAGCACAAATTTCAAGAGAGCTCAAGAA	723
Db			
Db	278	GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluLeuLysGlu	297
QY	724	TCCTTCAGTGTGTGAGAAATCTGGGTTCTTGAAGCTGCCCTCGCAGAGGTAGA	783
Db			

Db	298	SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg	317
QY	784	CTGGTGAATGCTACTTTTGGATATCTGGGATCATCGGACCCAGCTCAGCATCGCAAGTCA	843
Db			
Db	318	LeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla	337
QY	844	AGGATATGATGGGCAAGCTCAACCTCTGATTCAGGTGATCGATGATATTTATGATGTC	903
Db			
Db	338	ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspIleTyrAspVal	357
QY	904	TATGGCACCTTAGAAGAACTCGAACAATTCATCTGACCTCATTCGAAGATGGGATATAAC	963
Db			
Db	358	TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgTrpAspIleAsn	377
QY	964	TCAATCGACCAACTTCCGATTCACATGCAACTGTCTTCTTGCACTCAACAACTTCGTC	1023
Db			
Db	378	SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal	397
QY	1024	GATGATACATCGTACGATGTTATGAGAGGAGAAAGCGCTCAACGTTATATCCCTACCTCGG	1083
Db			
Db	398	AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg	417
QY	1084	CAATCGTGGGTTCATTTGGCGGATAAGTATATGTTAGAGGACCGGTGCTTCTACGGCGG	1143
Db			
Db	418	GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly	437
QY	1144	CACAAACCAAGTTTGAAGAGTATTGGAGAACTCATGGCAGTCGATAGTGGGCGCTGT	1203
Db			
Db	438	HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys	457
QY	1204	ATGTTAACGACACATATTTCCGAGTACACAGATTTCGTTCAAAAGGAGACCGTCGACAGT	1263
Db			
Db	458	MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer	477
QY	1264	TTCTACAAATCCACGATTTAGTTCGTTGGTCATCTTCGTCGCGCTTGTGCTGATGAT	1323
Db			
Db	478	LeuTyrLysTyrHisAspLeuValArgTrpSerPheValLeuLeuArgLeuAlaAspAsp	497
QY	1324	TTCGGAACCTCGTGGACAGAGTGACAGAGGGATGTCCGAAATCACTTCAGTGTCTAC	1383
Db			
Db	498	LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr	517
QY	1384	ATCAGTGACTACAACTCATCGGAGCGGAGCGGAGGACGTCGAAATGCTGATACGC	1443
Db			
Db	518	MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla	537
QY	1444	GAGGTGTGAAGAAAGTGAATCGGAGAGGGTGTGCAAGGATTTCTCATTCGCGCAAGAT	1503
Db			
Db	538	GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp	557
QY	1504	TTTATAGGATGTGCACTTTCATTTAGGAAGATGGCGCAGTTGATGTACCATTAATCGGAT	1563
Db			
Db	558	PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp	577
QY	1564	GGGACGGCACACACACCCCTATTATACATCAACAAATGACAGAACCTTATTCGAGCCC	1623
Db			
Db	578	GlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGluPro	597
QY	1624	TTTGCA 1629	
Db			
Db	598	PheAla 599	

RESULT 2
 Q9SW76 PRELIMINARY; PRT; 599 AA.
 ID Q9SW76
 AC Q9SW76;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Limonene synthase.
 OS Mentha longifolia (Horse mint).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;

OC lamids; Lamiaceae; Nepetoideae; Menthae; Mentha.
 OX NCBI_TaxID=38859;
 RN [1]
 RA Crock J., Coreau R.;
 RT "limonene synthase cDNA from Mentha candicans";
 RL Submitted (Aug-1999) to the EMBL/GenBank/DBSJ databases.
 DR EMBL; AF175323; AAD50304.1; -;
 DR HSP; Q40577; SEAU.
 DR GO: 0016829; F: lyase activity; IEA.
 DR GO: 0008152; P: metabolism; IEA.
 DR InterPro: IPR005630; Terpenoid synth. C.
 DR InterPro: IPR008949; Terpenoid synth.
 DR InterPro: IPR008930; Terp. cyc.teroid.
 DR InterPro: IPR001906; Terp. synth.-like.
 DR Pfam: PF01357; Terpene synth. 1.
 DR Pfam: PF03936; Terpene synth. C; 1.
 SQ SEQUENCE 599 AA; 69905 MW; BC789A0DD7D894EE CRC64;

Alignment Scores:
 Pred. No.: 5 46e-215 Length: 599
 Score: 2839.00 Matches: 534
 Percent Similarity: 98.89% Conservative: 2
 Best Local Similarity: 98.52% Mismatches: 6
 Query Match: 97.63% Indels: 0
 DB: 10 Gaps: 0

US-09-938-956-6 (1-1632) x Q9SW76 (1-599)

QY 4 AGAGCTCCGGAACCTACACCTCTCTCGTGGGTGTCACACTTCATCCATCCACTCTCTC 63
 DB 58 ArgArgSerGlyAsnTrpAsnProSerArgTrpAspValGluPheIleGlnSerLeuHis 77
 QY 64 AGTGACTATAAGGAGCAACAAACGTTAGGGTCTCTGAGCTCGTCACTTTGGTGAAG 123
 DB 78 SerAspTyrGluGluAspLysHisAlaIleArgAlaSerGluLeuValThrLeuValLys 97
 QY 124 ATGGAACTGAGAAAGAAAGCGATCAATTCCGACAACTTGATGTCATCGATCGACTTCGAG 183
 DB 98 MetGluLeuGlnLysGluThrAspHisIleArgGlnLeuGluLeuIleAspAspLeuGln 117
 QY 184 AGGATGGGGTGTCCGATCATTTCCAAATAGTTTCAAGAAATCTCTCTCTATATAT 243
 DB 118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIleTyr 137
 QY 244 CTGACCATCATATTACAGAACCTTTTCCAAAGAGAGAGAGGATCTCTACTCCACA 303
 DB 138 LeuAspHisHisTyrTrpLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157
 QY 304 TCTCTTGCAATTTAGGCTCTCTCAGAGAACATGGTTTTCAAGTCGCCACAGAGGATTCGAT 363
 DB 158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177
 QY 364 AGTTTCAGACGAGAGGGTGTGATTCAAAGAAAGCCTTAGCGAGCACACAGAGGATG 423
 DB 178 SerPheLysAsnGluGluGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197
 QY 424 TTGCAACTGATGAAGCTCTCTTCTGTGTGACGAGCGAAACCCAGCTCGAGTCAGCG 483
 DB 198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217
 QY 484 AGGAAATCCGACCAATTTTGGAGGAAAGTCAAGCGGGTGTGTGATGGCGAC 543
 DB 218 ArgGluPheAlaThrLysPheLeuGluGluArgValAsnGluGlyGlyValAspGlyAsp 237
 QY 544 CTTTAAACAGAAATCGCATATCTTTTGGACATCCCTCTTCATTTGGAGGATTAAGGCCA 603
 DB 238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTyrTrpArgIleLysArgPro 257
 QY 604 AATGCACTGTGTGGATCGAATGATAGAGAGCGCCGACATCAATCCAGTACTGTG 663
 DB 258 AsnAlaProAlaTrpIleGluTrpTyrArgLysAsnProAspMetAsnProValValLeu 277

QY 664 GAGCTTGCCTACTCGACTTAATATTTGTCACGACAAATTTCAAGAGAGCTCAAGAA 723
 DB 278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluGluLeuLysGlu 297
 QY 724 TCCTTCAGGTGCTGAGAAATCTGGTTTCTTGAGAGCTGCCCTTCGCAAGGATAGA 783
 DB 298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317
 QY 784 CTGGTGAATGCTACTTTTGGAAATCTGGGATCATCGAGCCACCTCAGCATGCAAGTCA 843
 DB 318 LeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla 337
 QY 844 AGSATAATGATGGCAAAAGTCAACGCTCTGATTCGGTGTGATCGATGATATTTATGATGC 903
 DB 338 ArgIleMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357
 QY 904 TATGGACCTTAGAAGAACTCGAACAAATCTACTGACTCATTCGAAAGATGGGATATAAC 963
 DB 358 TyrGlyThrLeuGluGluLeuGlnPheThrAspLeuIleArgTrpAspIleAsn 377
 QY 964 TCAATCGACCAACTTCCGATTTACATGCAACTGTGCTTTCTTGCACCACTTCGTC 1023
 DB 378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397
 QY 1024 GATGATACATCGTACGANTGTTATGAAGGAGAAAGCGTCAACGTTATACCTACTCGCGG 1083
 DB 398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417
 QY 1084 CAATCGTGGTGTGATTTGGCGGATAAGTATATGTTAGAGGCACGCTGGTCTTACGGCGG 1143
 DB 418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437
 QY 1144 CACAAACCAAGTTTGGAGAGATTTGGAGAACTCATCGGAGTGGATGAAGTGGCGCTGT 1203
 DB 438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457
 QY 1204 ATGTTAAACCAATATCTTCCGAGTAAAGATTCGTTTCAAAAGGAGACCGTCAAGCT 1263
 DB 458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477
 QY 1264 TTGTACAAATACCAAGATTTAGTTGCTGGTGTGATCTCTGCTTCTGGCGCTTGCTGATGAT 1323
 DB 478 LeuTyrLysTyrHisAspLeuValArgTrpSerPheValLeuArgLeuAlaAspAsp 497
 QY 1324 TTGGGAACCTCGGTGGAAGAGTGAAGAGCGGATGTCGGAATCACTTCAGTGTCTAC 1383
 DB 498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517
 QY 1384 ATGAGTGACTACAATGCAATCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 1443
 DB 518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla 537
 QY 1444 GAGGTGTGGAAGAAAGATGAATGCGGAGAGCGGTGTCGAGAGGATTTCTCCATTCGCAAGAT 1503
 DB 538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557
 QY 1504 TTATAGATGTGCAAGTTGATTTAGGAAGGATGGCGGAGTGTGATGTGATGATGATGATG 1563
 DB 558 PheIleGlyCysAlaAlaAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577
 QY 1564 GCGCAGCGGCACACACCTTAT 1623
 DB 578 GlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGluPro 597
 QY 1624 TTGGA 1629
 DB 598 PheAla 599

RESULT 3

Q940E7 PRELIMINARY; PRT; 613 AA.
 ID Q940E7;
 AC Q940E7;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE D-limonene synthase.
 OS Agastache rugosa.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Lamiales; Lamiaceae; Nepetoideae; Menthae; Agastache.
 OC NCBI_TaxID=39271;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Saeki D., Maruyama T., Ito M., Honda G.;
 RT "Cloning and in vitro expression of d-limonene synthase from Agastache
 rugosa";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY055214; AAL17636.1; - - - - -
 DR GO: GO:0016829; P:lyase activity; IEA.
 DR GO: GO:0008152; P:metabolism; IEA.
 DR InterPro: IPR005630; Terpenes synthase C.
 DR InterPro: IPR008949; Terpenoid synthase.
 DR InterPro: IPR008930; Terpene synthase.
 DR InterPro: IPR001906; Terpene synthase-like.
 DR Pfam: PF01397; Terpene synthase 1.
 DR Pfam: PF03936; Terpene synthase C; 1.
 SQ SEQUENCE 613 AA; 71422 MW; A92502C399P47572 CRC64;

Alignment Scores:
 Pred. No.: 6,49e-163 Length: 613
 Score: 2178.00 Matches: 407
 Percent Similarity: 86.40% Conservative: 63
 Best Local Similarity: 74.82% Mismatches: 72
 Query Match: 74.90% Indels: 2
 DB: 10 Gaps: 2

US-09-938-956-6 (1-1632) x Q940E7 (1-613)

QY 4 AGACGATCGGAATCAACCCCTTCGTTGGAGTCAACTTCATCCAACTCGCTTC 63
 DB 70 ArgArgSerGlyAsnTyrSerProSerArgTyrAspValAspPheIleGlnSerLeuAsn 89
 QY 64 AGTCACTATAGGAGGCAACACGATGATTAGGCTTCGAGCTGCTGCTGCTGAG 123
 DB 90 SerAspTyrGlnGluGluArgHisThrArgArgAlaSerGlnLeuIleThrGlnValLys 109
 QY 124 ATGGAACCTGGAGAAGAA--ACGGATCAAAATCGACAACTTGAGTTGATCGATCTG 180
 DB 110 MetLeuMetGluLysGluThrThrAspProIleArgGlnLeuLeuIleAspAspLeu 129
 QY 181 CAGAGGATGGGCTGCTCGATCATTTCCAAATGAGTTCAAAGAAATCTGCTCTATA 240
 DB 130 GlnArgLeuGlyLeuSerAspHisPheGlnAsnGluPheLysGlnLeuLeuAsnThrIle 149
 QY 241 TATCTGACCATCACTATTACAAG--AACCCCTTTCCAAAGAAAGAGGATCTCTAC 297
 DB 150 TyrLeuAspAsnLysTyrTyrAsnIleAsnIleMetArgGluGluSerArgAspLeuTyr 169
 QY 298 TCCACATCTTGGATTTAGCTTCTCCTCAGACACATGTTTCAAGTCCCAAGAGTA 357
 DB 170 SerThrAlaLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluVal 189
 QY 358 TTCGATAGTTTCAAGAACGAGGAGGTGAGTTCAAAGAAAGCCTTAGCGACACACAGA 417
 DB 190 PheGluCysPheLysAsnGluGluGlyAspPheLysAlaSerLeuIleAspAspThrArg 209
 QY 418 GGATTTGTCACACTGTATGAGTCTCTCTCTGTCGCGAAGCGCAACACCGCTCGAG 477
 DB 210 GlyLeuLeuGlnLeuTyrGluAlaSerPheLeuPheLysGluGlnLeuAsnThrLeuGlu 229
 QY 478 TCAGCGAGGGAATTCGCCACCAATTTTGGAGAAAGTGAACGAGGCTGCTGTGAT 537
 DB 230 IleAlaArgGluPheThrThrLysIleLeuGlnGluLysLeuLysGlyAspGluIleAsp 249
 QY 538 GCGACCTTTTAAACAGATCGCATATTTTGGACATCCCTCTTCAATGGAGGATTAA 597

Db 250 AspAsnLeuLeuSerSerIleArgTyrSerLeuGluIleProAsnTyrTrpSerValVal 269
 QY 598 AGGCCAAATGCACCTGTGTGATCGAATCGTATAGGAAGAGGCGCCACATGAATCCAGTA 657
 Db 270 ArgProAsnValSerValTrpIleAspGluTyrArgLysArgSerAspMetAsnProVal 289
 QY 658 GTGTTGGAGCTTCCCATCTCGACTTAATATTTTCAAGCACAATTTCAAGAGGCTC 717
 Db 290 ValLeuGluLeuAlaIleLeuAspAlaAsnIleValGlnAlaGlnLeuGlnLeu 309
 QY 718 AAGAATCTTCAGGTGGTGGAGAAATACTGGGTTTGTGAGAAGCTGCCCTTCGCAAG 777
 Db 310 LysGluSerLeuArgTyrTrpArgAsnThrCysPheValGluLysLeuProPheAlaArg 329
 QY 778 GATAGACTGTGGAATGCTACTTTTGGAAATCTGGCATCTGAGCCACTGAGCCACTGCA 837
 Db 330 AspArgLeuIleGluSerTyrPheTrpSerThrGlyMetValGluProArgGlnHisAla 349
 QY 838 AGTCAGAGGATAATGATGGGCAAGTCACGCTCTGATTACGGTGCATGATGATTTAT 897
 Db 350 AsnAlaArgIleIleMetAlaLysValIleAlaLeuIleThrValMetAspIleTyr 369
 QY 898 GATGCTCTATGGACCTTAGNAGACTCGAACATTCACCTGACCTCATTCGAGAGTGGAT 957
 Db 370 AspValTyrGlyThrLeuGluGluGluGlnPheThrGluAlaPheArgTrpAsp 389
 QY 958 ATAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTCACACTCAACAC 1017
 Db 390 ValSerSerIleAspGlnLeuProThrTyrMetGlnLeuCysPheLeuAlaIleAsnAsn 409
 QY 1018 TTGCTGATGATCATCGTACGATGATGAGGAGGAAAGGGTCAACGTTATACCCCTAC 1077
 Db 410 PheValAspAspThrAlaTyrAsnValLeuLysGluSerGlyValAsnValMetThrTyr 429
 QY 1078 CTGCGCAATCTGGGTTGATTTGGCGGATAAGTATATGATGAGAGCACGCTGGTCTTAC 1137
 Db 430 LeuArgLysSerTrpValAspGlnAlaGluAsnTyrLeuMetGluSerLysTrpTyr 449
 QY 1138 GCGGCGCACAAACCAAGTTTGGAGAGTATTTGGAGACTCATGTCAGTCGAGTATAGTGG 1197
 Db 450 SerGlyHisLysProSerLeuGluGluTyrLeuGluAsnSerTrpIleSerValSerGly 469
 QY 1198 CCTGTATGTTAAACGACATATTTCTCCAGTAAACAGATTCGTTCAAAAGAGAGCCGTC 1257
 Db 470 ProCysValLeuThrHisGluPheGlyValThrAspSerLeuAlaLysAspThrLeu 489
 QY 1258 GACAGTTTGTACAAATACCAACGATTTAGTTCGTTGCTCATCTCTGCTGCGGCTTCT 1317
 Db 490 AspSerLeuTyrGluTyrHisAspIleValArgTrpSerSerTyrLeuLeuArgLeuAla 509
 QY 1318 GATGATTTGGAACTCGGTGGAGAGGTGACGACAGGGGATGTCCGAAATCACTTCAG 1377
 Db 510 AspAspLeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerIleGln 529
 QY 1378 TGCTACATGATGATCATCATCGAGGCGGAGCGCGGAGCAACGTCGTAATGGCTG 1437
 Db 530 CysTyrMetHisAspAsnAspAlaSerGluGluAlaArgGlnHisIleLysGlyLeu 549
 QY 1438 ATAGCGAGGTGTGGAAGAGATGAATCGCGAGAGGGTGTGGAAGATTTCCATTCGCG 1497
 Db 550 IleArgGluMetTrpLysLysMetAsnValGluArgValSerGluAspSerProPheCys 569
 QY 1498 AAGATTTTATAGGATGCTGATTTAGCAAGATGGCGGAGTGTGATGTATACCATAT 1557
 Db 570 ArgAspPheIleArgCysGluAspLeuGlyArgMetAlaGlnPheMetTyrHisTyr 589
 QY 1558 GGAGATGGGCAACGACACACACCTATTATATATCAATCAACAAATGACCAAGACCTTATTC 1617
 Db 590 GlyAspGlyHisGlyThrGlnHisProLysIleHisGlnGlnIleAlaAlaCysLeuPhe 609
 QY 1618 GAGCCCTTTGCA 1629
 Db 610 GlnProPheAla 613

RESULT 4
Q9FWU5
ID Q9FWU5 PRELIMINARY; PRT; 620 AA.
AC Q9FWU5;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE (+)-4R-limonene synthase.
OS Schizonepeta tenuifolia.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Lamiaceae; Nepetoideae; Menthaceae; Schizonepeta.
OX NCBI_TaxID=135200;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2120190; PubMed=11305598;
RA Maryama T., Ito M., Kluchi F., Honda G.;
RT "Molecular cloning, functional expression and characterization of d-
limonene synthase from Schizonepeta tenuifolia.";
RL Biol. Pharm. Bull. 24:373-377(2001).
DR EMBL; AF282875; AAC01140.2; -;
DR HSP; O40577; SEAU.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005630; Terpene synth. C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp. cyc. tocid.
DR InterPro; IPR001906; Terp. synth.-like.
DR Pfam; PF01397; Terpene synth. 1.
DR Pfam; PF03936; Terpene synth. C.1.
SQ SEQUENCE 620 AA; 71810 MW; 9AB8E7302442C55E CRC64;

Alignment Scores:
Pred. No.: 2,896-158 Length: 620
Score: 2119.00 Matches: 401
Percent Similarity: 84.31% Conservative: 61
Best Local Similarity: 73.18% Mismatches: 80
Query Match: 72.8% Indels: 6
DB: 10 Gaps: 3

US-09-938-956-6 (1-1632) x Q9FWU5 (1-620)

QY	4	ACAGCATCGGAAATACAAACCTTCCTGGTGGATGTCACATTCATCCATCGCTTC	63
DB	73	ArgArgSerGlyAsnTyrLysProSerArgTrpAspValAspPheMetGlnSerLeuAsn	92
QY	64	ACTGACTATAAGGAGGACAAACACGTGATTAGGCTTCAGCTGGTGCATCTTGGTGAAG	123
DB	93	SerAspTyrGlnGluGluArgHisArgThrLysAlaSerGluLeuLeuThrGlnVallys	112
QY	124	ATGGAACTGGGAAGAAAGC-----GATCAAAATCGCAACTTGAGTTGATCGATGAC	177
DB	113	AsnLeuLeuGluLysGluThrSerAspAspProIleArgGlnLeuGluLeuLeuAspAsp	132
QY	178	TTGCGAGAGATGGGGCTGCCATCATTTCCAAAATGAGTTCAAGAAATCTTGCTCTCT	237
DB	133	LeuGlnArgLeuGlyLeuSerAspHisPheGluHisGluPheLysGluValLeuAsnSer	152
QY	238	ATATATCTGCACATCATTTATCAAGAACCCCTTTTCCAAAGAA-----GAAAGG	288
DB	153	IleTyrLeuAspAsnLysTyrTyrAsnIleAsnIleMetLysGluThrThrSerSerArg	172
QY	289	GATCTCTACTCCACATCTCTTCATTTAGGCTTCCTCAGAGACATGGTTTTCAGTCGCA	348
DB	173	AspLeuTyrSerThrAlaLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValala	192
QY	349	CAAGAGGTATTCATAGTTTCAAGAACGAGGAGGTGAGTTCAAGAAAGCCCTTAGCGAC	408
DB	193	GlnGluValPheAspCysPheLysAsnGluGluGlyGluPheLysAlaSerLeuSerAsp	212
QY	409	GACACAGAGGATTTGTCACATGATGAGCTTCCTTCTCTTCGACGGAAGCGCAACC	468

DB	213	AspProArgGlyLeuLeuGlnLeuTyrGluAlaSerPheLeuPheLysGluGlyGluAsn	232
QY	469	ACGCTCGAGTCAGCGAGGGAATTCGCCACCAAAATTTTCGAGGAGAAAGTGAACGAGGT	528
DB	233	ThrLeuGluIleAlaArgGluPheAlaThrLysLeuLeuGlnGluLysValAsnSerSer	252
QY	529	GGT---GTTGATGGGACCTTTTAAACAAGATCGCATATTCTTTGACACATCCCTCTCAT	585
DB	253	AspGluIleAspAsnLeuLeuSerSerIleArgTyrSerLeuGluIleProThrTyr	272
QY	586	TGGAGATTAAAGGCGCAATCACCTGTGTGGATCGATGATATAGGAGAGGCCCGCAC	645
DB	273	TrpSerValIleArgProAsnValSerValTrpIleAspAlaTyrArgLysArgProAsp	292
QY	646	ATGAATCCAGTAGTGTGGAGTTCGCATCTCGACTTAATATATCTTCAAGCACAAATTT	705
DB	293	MetAsnProValValLeuGluLeuAlaIleLeuAspAlaAsnIleMetGlnAlaGlnLeu	312
QY	706	CAAGAGAGCTCAAGAATCCCTTCAGGTGGTGGAGAAATCTGGTTCCTGAGAGCTG	765
DB	313	GlnGlnGluLeuLysGluAlaLeuGlyTrpTrpArgAsnThrTrpPheValGluLysLeu	332
QY	766	CCCTTCGCAAGGATAGACTGGTGGAAATCTACTTTTGGAAATCTGGGATCATCGGCCA	825
DB	333	ProPheAlaArgAspArgLeuValGluSerTyrPheTrpSerThrGlyMetValProArg	352
QY	826	CGTCAGCATGCAAGTCAAGGATATGATGGGCAAGTCAACGCCTGATACGGTGATC	885
DB	353	ArgGlnHisLysThrAlaArgLeuLeuMetAlaLysValIleAlaLeuIleThrValMet	372
QY	886	GATGATATTATGATGCTATGGCACTTAGAAGAACTCGAACAAATCTACTGACCTCAT	945
DB	373	AspAspIleTyrAspValTyrGlyThrLeuGluLeuGluLeuPheThrAspAlaPhe	392
QY	946	CGAAGATGGGATATAAACTCAATCGACCAACTTCGCGATACATGCAACTGCTTCCT	1005
DB	393	ArgArgTrpAspValSerSerIleAspHisLeuProThrTyrMetGlnLeuCysPheLeu	412
QY	1006	GCACTCAACAACTTCGTCGATCATCATCGTACGATGTTATGAAGGAGAAAGCGCTCAAC	1065
DB	413	SerIleAsnAsnPheValValAspThrAlaTyrAsnIleLeuLysGluThrGlyValAsn	432
QY	1066	GTATATACCTTACCTCGGCAATCGTGGTGTGATTTGGCGGATAGTATATGTTAGAGCA	1125
DB	433	ValThrThrTyrLeuGluLysSerTrpValAspGlnAlaGluAsnTyrLeuMetGluSer	452
QY	1126	CGTGGCTTCTACGGCGGCGACAAACCAAGTTGGAAGATGATTGGAGAACTCATGGCAG	1185
DB	453	LysTrpPheTyrSerGlyHisLysProSerLeuAspGluTyrLeuGluAsnSerTrpIle	472
QY	1186	TCGATAAGTGGCCCTGTATGTTTAAACGACATATTTCCGAGTAAACAGATTCTGCACA	1245
DB	473	SerValSerGlyProCysValLeuThrHisGluPhePheGlyValThrAspSerLeuAla	492
QY	1246	AAGAGACCTTCACAGATTGTGTAACAATACCAGATTAGTTCGTTGGTTCATCTTCGTT	1305
DB	493	LysAspThrLeuAspSerLeuTyrGluTyrHisAspIleValArgTrpSerSerTyrLeu	512
QY	1306	CTGGGCTTCCTCATATTGGGAACTCGGTGGAAGAGGTGAGCAGAGGGATGTGCCG	1365
DB	513	LeuArgLeuAlaAspAspLeuGlyThrSerValGluGluValSerArgGlyAspValPro	532
QY	1366	AAATCACTTCAGTGTACATGACTGACTACAAATCGATCGGAGCGGCGGAGGAGCAC	1425
DB	533	LysSerIleGlnCysTyrMetAsnAspAsnAlaSerGluGluAlaArgGluHis	552
QY	1426	GTGAATGGCTGATAGCGGAGGTGTGGAAGAAGATGAATCGGAGAGGGGTGTGCAAGGAT	1485
DB	553	ValLysGlyLeuIleArgValMetTrpLysLysMetAsnAlaGluArgValSerGluAsp	572
QY	1486	TCTCCATTCGGCAAGATTTTATGAGATGTCAGTTGATTTAGAGAGGATGGCGAGTTC	1545
DB	573	SerProPheCysLysAspPheIleArgCysCysGluAspLeuGlyArgMetAlaGlnPhe	592


```

QY 1423 CACGTGAATGGCTGATGCGGAGGTGTGGAGAGAGATGAATGCGGAGGGTGTGGAAG 1482
DB 536 HisValArgTTPLeuIleAlaGluThrTrpLysLysIleAsnGluGluValTrpSerAla 555
QY 1493 GATTCTCAATCGCGAAGATTTTATAGGATGTCAGTGTGATTTAGCAAGATGCGGCGAG 1542
DB 556 AspSerProPheCysLysAspPheIleAlaCysAlaAlaAspMetGlyArgMetAlaGln 575
QY 1543 TTGATGTACCAATATGAGATGGGCGGCGGACACACACACCCCTATTATACATCAACAATG 1602
DB 576 PheMetTyrHisAsnGlyAspGlyHisGlyIleGlnAsnProGlnIleHisGlnGlnMet 595
QY 1603 ACCAGAACCTTATTCGAG 1620
DB 596 ThrAspIleLeuPheGlu 601
RESULT 6
ID Q9FV75 PRELIMINARY; PRT; 604 AA.
AC Q9FV75;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Limonene synthase.
OS Perilla citriodora.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Lamiaceae; Nepetoideae; Elaeagnaceae; Perilla.
OX NCBI_TaxID=121094;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=5601;
RA Dodo R., Ito M., Takeya H., Honda G.;
RT "Limonene synthase from Perilla sp.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF241790; AAG31435.1; -.
DR HSP; O40577; SEAT.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005630; Terpene synth. C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp. cyc. toroid.
DR InterPro; IPR001906; Terp. synth.-like.
DR Pfam; PF01397; Terpene synth. 1.
DR Pfam; PF03936; Terpene synth. C. 1.
SQ SEQUENCE 604 AA; 70498 MW; 60363DE340721A1C CRC64;

Alignment Scores:
Pred. No.: 2,29e-149 Length: 604
Score: 2006.00 Matches: 375
Percent Similarity: 82.23% Conservative: 74
Best Local Similarity: 68.68% Mismatches: 85
Query Match: 68.98% Indels: 12
DB: 10 Gaps: 3

US-09-938-956-6 (1-1632) x Q9FV75 (1-604)
QY 4 AGCAGATCGGAACTACAAACCTTCTCGTTGGGATGTCACCTTCATCCAACTCGTTCTC 63
DB 61 ArgArgSerGlyAsnTyrSerProSerPheTrpAsnAlaAspTyrIleLeuSerLeuAsn 80
QY 64 AGTGAATTAAGAGGAC---AAACAGCTGATAGGGTCTTCAGCTGGTCACCTTGGTG 120
DB 81 AsnHisTyrLysGluGluSerArgHisMetLysArgAlaGlyGluLeuIleValGlnVal 100
QY 121 AAGATGGAACCTGGAGAAAGACGATCAAAATCCGACAACTTGAGTTGATCGATGACTTG 180
DB 101 LysMetValMetGlyLysGluThrAspProValValGlnLeuGluLeuIleAspAspLeu 120
QY 181 CAGAGGATGGGCTGTCCGATCATTTCCAAATATGATTTCAAGAATCTTGTCTCTATA 240
DB 121 HisLysLeuAlaLeuSerHisHisPheGluLysGluIleLysGluIleLysPheAsnIle 140

```

```

QY 241 TATCTCGACCATCACTATTATTACAAGAACCCCTTTTCAAAAAGAAAGGATCTTACTCC 300
DB 141 SerIleTyrAspHisLysIle-----MetValGluArgAspLeuTyrSer 155
QY 301 ACATCTCTTCGATTTAGGCTCTCAGAGAACATGGTTTCAAGTCGCACAAAGAGGTATTC 360
DB 156 ThrAlaLeuAlaPheArgLeuLeuArgGlnTyrGlyPheLysValProGlnGluValPhe 175
QY 361 GATAGTTTCAAGAACGAGGAGGTGAGTTTCAAGAAAGCCTTAGCGACACACAGAGA 420
DB 176 AspCysPheLysAsnAspAsnGlyGluPheLysArgSerLeuSerSerAspThrLysGly 195
QY 421 TTGTTGCAACTGTATGAAGCTTCTCTTCTGTTGAGGAGGCGGAAACACGCTCCAGTCA 480
DB 196 LeuLeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluMetThrLeuGluLeu 215
QY 481 GCGAGGGAATTCGCCACCAAAATTTTGGAGGAAAAAGTCAACGAGGGGTGTGTGATGCC 540
DB 216 AlaArgGluPheAlaThrIlePheLeuGlnGluLysLeuAsnAspLysThrIleAspAsp 235
QY 541 GAC-----CTTTTACAAAGATCGCATATTTCTTTGACACATCCCTCTT 582
DB 236 AspAspAlaAspThrAsnLeuIleSerCysValArgHisSerLeuAspIleProfile 255
QY 583 CATTGAGGATTAAGGCGCAATGCACCTGTGTGATCGAATGTATAGGAGAGGCC 642
DB 256 HisTrpArgIleGlnArgProAsnAlaSerTrpTrpIleAspAlaTyrLysArgSer 275
QY 643 GACATGAATCCAGTAGTGTGTGAGCTTGCATCTGCTAAATATTCTTCAAGCACA 702
DB 276 HisMetAsnProLeuValLeuGluLeuAlaLysLeuAspLeuAsnIlePheGlnAlaGln 295
QY 703 TTTCAAGAGAGCTCAAGAAATCCITCAGTGTGTGAGAGAAATCTGGGTTTGTGAGAAG 762
DB 296 PheGlnGlnGluLeuLysGlnAspLeuGlyTrpTrpLysAsnThrCysLeuAlaGluLys 315
QY 763 CTGCCCTTCGCAAGGATAGACTGTGTGGAATGCTACTTTTGGAAATCTCGGATCATCGAG 822
DB 316 LeuProPheAlaArgAspArgLeuValGluCysTyrPheTrpCysThrGlyIleLeuGln 335
QY 823 CACGTCAGATCGCAAGTCAGGATAATGATGGGCAAGTCAACGCTGTGATAGGNG 882
DB 336 ProLeuGlnHisGluAsnAlaArgValThrLeuAlaLysValAsnAlaLeuIleThr 355
QY 883 ATCGATGATATTATGATGCTCTATGCGACCTTAGAAGAACTCGAACTACTGACTC 942
DB 356 LeuAspAspIleTyrAspValTyrGlyThrLeuGluGluLeuPheThrGluAla 375
QY 943 ATTGGAAGATGGATATAAATCAATCGACCACTCCGATTCAGTCAACTGTCTTT 1002
DB 376 IleArgArgTrpAspValSerSerIleAspHisLeuProAsnTyrMetGlnLeuCysPhe 395
QY 1003 CTTGCACTCAACAACTTCGTGCGATGATACATCGTACGATGTTATGAGGAGAAAGCGTC 1062
DB 396 LeuAlaLeuAsnAsnPheValAspAspThrAlaTyrAspValMetLysGluLysAspIle 415
QY 1063 AACGTTATACCTTACTCGGCAATCGTGGTGTGATTTGGCGATAGTATATGATAGAG 1122
DB 416 AsnIleIleProTyrLeuArgLysSerTrpLeuAspLeuAlaGluThrTyrLeuValGlu 435
QY 1123 GCAGGTGTGTTCTACGGCGGCAACCAAGTTTCGAAGAGTATTGGGAACACTCATGG 1182
DB 436 AlaLysTrpPheTyrSerGlyHisLysProAsnMetGluGluTyrLeuAsnAsnAlaTrp 455
QY 1183 CAGTCGATAGTGGGCCCTGTATGTTTAACGCACATATTTCTCGAGTACAGATTCGTC 1242
DB 456 IleSerIleSerGlyProValMetLeuCysHisValPhePheArgValThrAspSerIle 475
QY 1243 ACAAGGAGCCCTCCAGATTTGTCAAAATACACGATTTAGTTCTGTTGGTTCATCTTC 1302
DB 476 ThrArgGluThrValGluSerLeuPheLysTyrHisAspLeuIleArgTyrSerSerThr 495
QY 1303 GTTCTCGCGCTTCTGCTGATGATTGTTGGAACTCGTGGTGAAGAGGTGAGAGGGGATGTG 1362

```

```

Db 496 ILeuAArgLeuAlaAspLeuGlyThrSerLeuGluGluValSerArgGlyAspVal 515
QY 1363 CGGAATCAGTTCAGTCACTAGTCACTACATCATCGAGGCGGAGGCGGGAAG 1422
Db 516 ProLysSerIleGlnCysTyrMetAsnAspAsnAlaSerGluGluAlaArgArg 535
QY 1423 CAGTGAAATCGCTAGTAGCGGAGGTGTGAAGAGATGAATCGGAGAGGGGTGCGAAG 1482
Db 536 HisValArgTrpLeuIleAlaGluThrTrpLysIleAsnGluGluValTrpSerAla 555
QY 1483 GATTCCTCATTCGGCAAGATTTATAGGATGTGCAGTTGATTTAGGAAGGTGGCGCAG 1542
Db 556 AspSerProPheCysLysAspPheIleAlaCysAlaAspMetGlyArgMetAlaGln 575
QY 1543 TTGATGTACCAATATGAGATGGCGGCGGACACACACACCCCTATTATACATCAACAAATG 1602
Db 576 PheMetTyrHisAsnGlyAspGlyHisGlyIleGlnAsnProGlnIleHisGlnGlnMet 595
QY 1603 ACCAGAACCTTATTCGAG 1620
Db 596 ThrAspIleLeuPheGlu 601

RESULT 7
Q9M5A4 PRELIMINARY; PRT; 604 AA.
AC Q9M5A4;
DT 01-OCT-2000 (TRENDEL. 15, Created)
DT 01-OCT-2000 (TRENDEL. 15, Last sequence update)
DT 01-OCT-2003 (TRENDEL. 25, Last annotation update)
DE Limonene synthase.
OS Perilla citriodora.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Lamiaceae; Nepetoideae; Elsholtziaceae; Perilla.
NCBI_TaxID=121094;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=20189430; PubMed=10726896;
RA Ito M., Kiuchi F., Yang L.H., Honda G.;
RT "Perilla citriodora from Taiwan and its phytochemical
RL Biol. Pharm. Bull. 23:359-362 (2000).
DR EMBL; AF233894; AAF65545.1; -.
DR HSSP; Q40577; 5EAT.
DR GO; GO:0016829; P:lyase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005630; Terpene synth C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp cyc toroid.
DR Pfam; PF01397; Terpene synth-like.
DR Pfam; PF03936; Terpene synth; 1.
DR PF03936; Terpene synth C; 1.
SQ SEQUENCE 604 AA; 70597 MW; 6A3046E7D3A4C129 CRC64;

Alignment Scores:
Pred. No.: 2,29e-149 Length: 604
Score: 2006.00 Matches: 375
Percent Similarity: 82.23% Conservative: 74
Best Local Similarity: 68.68% Mismatches: 85
Query Match: 68.98% Indels: 12
DB: 10 Gaps: 3

US-09-938-956-6 (1-1632) x Q9M5A4 (1-604)
QY 4 AGAGATCCGGAAGAACTCAACCCCTCTCGTTGGATGTCACTTCATCCATCGCTTC 63
Db 61 ArgArgSerGlyAsnTyrSerProSerPheTrpAsnAlaAspTyrIleLeuSerLeuAsn 80
QY 64 AGTGACTATAGGAGGAC---AAACAGTGTATAGGCTCTCTCAGCTGCTCATTGGTG 120
Db 81 AsnHisTyrLysGluGluSerArgHisMetLysArgAlaGlyGluLeuIleValGlnVal 100

```

```

QY 121 AAGATGGAAGTGGAGAAAGAAACGGATCAAAATTCGACAACTTTCAGTTGATCGACTTG 180
Db 101 LysMetValMetGlyLysGluThrAspProValValGlnLeuGluLeuIleAspLeu 120
QY 181 CAGAGGATGGGTGTCCTCCGATCATTTCCAAATAGTTCAAAGAAATCTTCTCTCTATA 240
Db 121 HisLysLeuAlaLeuSerHisHisPheGluLysGluIleLysGluLeuPheAsnIle 140
QY 241 TATCTCGACCATCAGTATTACAAAGAACCCCTTTTCCAAAAGAAAGAGGATCTTACTCC 300
Db 141 SerIleTyrAspHisLysIle-----MetValGluArgAspLeuTyrSer 155
QY 301 ACATCTCTTCATTTAGGCTCTCCAGAACATGGTTTTCAGTGCACAAAGAGGTATTC 360
Db 156 ThrAlaLeuAlaPheArgLeuLeuArgGlnTyrGlyPheLysValProGlnLeuValPhe 175
QY 361 GATAGTTTCAAGAACGAGGAGGCTGAGTTCAAAGAAAGCCTTAGCGACACACAGAGA 420
Db 176 AspCysPheLysAsnAspAengGlyGluPheLysArgSerLeuSerSerAspThrLysGly 195
QY 421 TTGTTGCAACTGTATGAAGCTTCTTCTTCGAGGAGGCGGAAACACGCTCGAGTCA 480
Db 196 LeuLeuGlnLeuTyrGluAlaSerPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 215
QY 481 GCGAGGAAATTCGCCACCAAAATTTTTCGAGGAAAGATGAACGAGGCTGGTTCATGCC 540
Db 216 AlaArgGluPheAlaThrIlePheLeuGlnGluLysLeuAsnAspLysThrIleAsp 235
QY 541 GAC-----CCTTTAAACAAGATCGCATATTTCTTGACATCTCTTCTGACATCTCT 582
Db 236 AspAspAspAlaAspThrAsnLeuIleSerCysValArgHisSerLeuAspIlePro 255
QY 583 CATTGGAGGATTAAGGCGCAATGACCTGTGTGTGATCGAATGTGTATAGGAGGCGCC 642
Db 256 HisTrpArgIleGlnArgProAsnAlaSerTrpTrpIleAspAlaTyrLysArgSer 275
QY 643 GACATGAATCCAGTGTGTGGAGCTTGCATCTGCTTAAATATTTCTCAACACAA 702
Db 276 HisMetAsnProLeuValLeuGluLeuAlaLysLeuAspLeuAlaIlePheGln 295
QY 703 TTTCAGAGAGTCAAGAAATCCTTCAGGTGGTGGAGAAATCTGAGGTTTCTGAGAG 762
Db 296 PheGlnGlnLeuLysGlnLeuAspLeuGlyTrpTrpLysAsnThrCysLeuAlaGly 315
QY 763 CTGCTCTCGCAAGGATAGCTGTTGAAATCTCTTTTGGAAATCTGAGGATCTCAG 822
Db 316 LeuProPheAlaArgAspArgLeuValGluCysTyrPheTrpCysThrGlyIleLeu 335
QY 823 CCAGCTCAGCATCAAGTCAAGGATATGATGGGCAAGTCAACGCTCTGATACGGTG 882
Db 336 ProLeuGlnHisGluAsnAlaArgValThrLeuAlaLysValAsnAlaLeuIleThr 355
QY 883 ATCGATGATATTTATGATGCTATGTCATGACCTTAGAAGAACTCGAACATTCACCTC 942
Db 356 LeuAspAspIleTyrAspValTyrGlyThrLeuGluGluLeuLeuPheThrGluAla 375
QY 943 ATTGGAAGTGGATATAAAGTCAATCGACCAACTTCGCGATATGATGATGCTGCTTT 1002
Db 376 IleArgArgTrpAspValSerIleAspHisLeuProAsnTyrMetGlnLeuCysPhe 395
QY 1003 CTTGCACTCAACAACTTCGTGATGATCATCATGATGATGATGATGATGATGATGATG 1062
Db 396 LeuAlaLeuAsnAsnPheValAspAspThrAlaTyrAspValMetLysGluLysAsp 415
QY 1063 AACGTTATACCTTCTGCGGCAATCGTGGTGTGATTTGGCGGATAGTATATGATGATG 1122
Db 416 AsnIleProTyrLeuArgLysSerTrpLeuAspLeuAlaGluThrTrpLeuValGlu 435
QY 1123 GCACGGTGGTTCACGCGCGGCAACCAAGTTTGGAGAGATATTTGGAGAGATCATGG 1182
Db 436 AlaIysTrpPheTyrSerGlyHisLysProAsnMetGluGluTyrLeuAsnAlaTrp 455
QY 1183 CAGTCGATAGTGGGCGCTGTATGTTAAACGCACATATTTCTCCGAGTAAACAGATCGTTC 1242

```

```

Db      456 ILeSerIleSerGlyProValMetLeuCyHisValPhePheArgValThrAspSerIle 475
QY      1243 ACAAGAGGACCGTCGACAGTTTGTACAAATACACAGATTAGTTCGTGGTTCATCCCTTC 1302
Db      476 ThrArgGluThrValGluSerLeuPheLysTyrHisAspLeuIleArgTyrSerThr 495
QY      1303 GTTCTGGCGGCTTGTGATGATTTGGGAACCTCGTGGAGAGAGGTGAGCAGAGGGATGTG 1362
Db      496 ILeuArgLeuAlaAspAspLeuGlyThrSerLeuGluGluValSerArgGlyAspVal 515
QY      1363 CCGAAATACATCTCAGTGTACATGAGTACTCAATGTCATCGAGCGAGCGGCGGGAAG 1422
Db      516 ProLysSerIleGlnCysTyrMetAsnAspAsnAlaSerGluGluAlaArgArg 535
QY      1423 CAGTGAAATGGCTGTAGCGGAGGTGTGGAGAGAGATGAATGCGGAGAGGGTGTGGAAG 1482
Db      536 HisValArgTrpLeuIleAlaGluThrTrpLysLysIleAsnGluGluValTrpSerAla 555
QY      1483 GATTCCTCCATTCGGCAAGATTTTATAGGATGTGCAGTTGATTTAGGAGAGGTGCGCAG 1542
Db      556 AspSerProPheCysLysAspPheIleAlaCysAlaAlaAspMetGlyArgMetAlaGln 575
QY      1543 TTGATGTACATATATCGAGATGGCGACGACCAACACCCCTATTATATACATCAACAATG 1602
Db      576 PheMetTyrHisAsnGlyAspGlyHisGlyIleGlnAsnProGlnIleHisGlnGlnMet 595
QY      1603 ACCAGAAACCTTATTTCGAG 1620
Db      596 ThrAspIleLeuPheGlu 601

RESULT 8
ID      004806 PRELIMINARY; PRT; 603 AA.
AC      004806;
CT      01-JUL-1997 (TrEMBLrel. 04, Created)
DT      01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DE      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      LIMONENE cyclase.
GN      GPFLC.
OS      Perilla frutescens.
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC      Lamiales; Lamiaceae; Nepetoideae; Elsholtzieae; Perilla.
OX      NCBI_TaxID=48386;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=NO.9;
RX      MEDLINE=96400360; PubMed=8806736;
RA      Yuba A., Yazaki K., Tabata M., Honda G., Croteau R.;
RT      "cDNA cloning, characterization, and functional expression of 4S-(-)-
RL      limonene synthase from Perilla frutescens."
RN      Arch. Biochem. Biophys. 332:280-287(1996).
RP      SEQUENCE FROM N.A.
RA      Teubouchi T., Hara M., Yazaki K., Honda G.;
RL      Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR      EMBL; D49368; BAA08367.1; -;
DR      EMBL; AB005744; BAA21629.1; -;
DR      HSP; Q40577; SEAU
DR      GO; GO:0016829; P:lyase activity; IEA.
DR      GO; GO:0008152; P:metabolism; IEA.
DR      InterPro; IPR005630; Terpenoid synth. C.
DR      InterPro; IPR008949; Terpenoid synth.
DR      InterPro; IPR008930; Terp cyc toroid.
DR      InterPro; IPR001906; Terp synth-like.
DR      Pfam; PF01397; Terpene synth. 1.
DR      Pfam; PF03936; Terpene synth. C. 1.
SQ      SEQUENCE 603 AA; 70321 MW; 67F41F0C6C871181 CRC64;

Alignment Scores:
Pred. No.: 3.61e-149 Length: 603
Score: 2003.50 Matches: 375

```

```

Percent Similarity: 81.47% Conservative: 69
Best Local Similarity: 68.81% Mismatches: 90
Query Match: 68.90% Indels: 11
DB: 10 Gaps: 2

US-09-938-956-6 (1-1632) x 004806 (1-603)

QY      4 AGACGATCGGAAATACAAACCTTCTCGTTGGGATGTCAACTTCATCAATCGCTTCTC 63
Db      61 ArgArgSerGlyAsnTyrSerProSerPheTrpAsnAlaAspTyrIleLeuSerLeuAen 80
QY      64 AGTGACTATAAGAGGAGACAAACAGCTGATTAGGGCTTTCAGCTGCTCCTTTGGTGAAG 123
Db      81 SerHisTyrLysAspLysSerHisMetLysArgAlaGlyLeuIleValGlnValLys 100
QY      124 ATGGAACCTCGAGAAAGAAACCGATCAAAATTCGACAACTTGAGTTGATCGATCGACTTCGAG 183
Db      101 MetValMetGlyLysGluThrAspProValGlnLeuGluLeuLeuLeuLeuLeuGln 120
QY      184 AGATCGGGCTCTCCGATCAITTCAAAATGATGTTCAAGAAATCTTCTCTATATATAT 243
Db      121 LysLeuAlaLeuSerHisHisValGluLysGluLysGluLysGluLysGluLysGluLys 140
QY      244 CTCGACCATCACTATTACAGAAACCTTTTCCAAAGAAAGAAAGAGGATCTCTACTCCACA 303
Db      141 ThrTyrAspHisLysIle-----MetValGluArgAspLeuTyrSerThr 155
QY      304 TCTCTTGGATTTAGGCTCTCTCAGAGAACATGGTTTTCAGTCGCACAGAGATTCGAT 363
Db      156 AlaLeuAlaPheArgLeuLeuArgGlnTyrGlyPheLysValProGlnGluValPheAsp 175
QY      364 AGTTTCAAGAACAGGAGGAGTTCAGTTCAAGAAAGACCTTACGACGACGACAGAGGATG 423
Db      176 CysPheLysAsnAspAsnGlyLysPheLysArgSerLeuSerSerAspThrLysGlyLeu 195
QY      424 TTGCAACTGTATGAAGCTCTCTTCTGTGAGGAGGCGAAACCGCTGAGTTCAGTCAGCG 483
Db      196 LeuGlnLeuTyrGluAlaSerPheLeuLeuLeuThrGluGlyMetThrLeuGluLeuAla 215
QY      484 AGGGAATTCGCCACCAAAATTTTGGAGCAAAAAGTGAACGAGGGTGTGTGTGATGGGAC 543
Db      216 ArgGluPheAlaThrLysSerLeuGlnGluLysLeuAsnGluLysThrIleAspAsp 235
QY      544 -----CTTTTAAAGATCCGATATCTTTGGACATCTTTGGACATCCCTTCAT 585
Db      236 AspAspAlaAspThrAsnLeuLeuLeuSerCysValArgHisSerLeuAspIleProIleHis 255
QY      586 TGGAGGATTAAGGCGCAATTCACCTGTGTGGATCGAATGGTATAGAGAGAGCGCCGAC 645
Db      256 TrpArgGlnArgProAsnAlaSerTrpTrpIleAspAlaTyrLysArgArgSerHis 275
QY      646 ATGATCCAGTAGTGTGGAGCTTGCATCTCCACTTAATATATTTGTCAGACCAATTT 705
Db      276 MetAsnProLeuValLeuLeuAlaLysLeuAspLeuAsnIlePheGlnAlaGlnPhe 295
QY      706 CAAGAAGAGCTCAAGATCCCTTCAGGTGTGGAGAAATACTGGGTTTGTGAGAACTG 765
Db      296 GlnGlnGluLeuLysGlnAspLeuGlyTyrTrpLysAsnThrCysLeuAlaGluLysLeu 315
QY      766 CCTTCGCAAGGATAGACTGTGGATGCTACTTTTGGATATCTGGATCATCGACCCA 825
Db      316 ProPheValArgAspArgLeuValGluCysTyrPheTrpCysThrGlyIleIleGlnPro 335
QY      826 CGTCAGCATGCAAGTGAAGGATATGATGGCGAAAGTCAACGCTCTGTGATTCAGGTATC 885
Db      336 LeuGlnHisGluAsnAlaArgValThrLeuAlaLysValAsnAlaLeuIleThrThrLeu 355
QY      886 GATGATATTATGATGCTATGGCAGCTTAGAGACTCGAAGACTCGAAGACTCAGCTCAT 945
Db      356 AspAspIleTyrAspValTyrGlyThrLeuGlnGluLeuLeuPheThrGluAlaIle 375
QY      946 CGAAGATGGATATAAACTCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTCTT 1005

```

Db 376 ArgArgTrpAspValSerSerIleAspHisLeuProAsnTyrMetGlnLeuCysPheLeu 395
 QY 1006 GCATCAACAACTTCGTCGATGATACATCGTACGATGTTATGAAGAGAGAAAGCGCTCAAC 1065
 Db 396 AlaLeuAsnAsnPheValAspAspThrAlaTyrAspValMetLysGluLysAspIleAsn 415
 QY 1066 GTTATACCTACCTCGCGCAATCGTGGTGTGATTTGGCGATTAAGTATATGTAGAGGCA 1125
 Db 416 IleIleProTyrLeuArgLysSerTrpLeuAspLeuAlaGluThrTyrLeuValGluAla 435
 QY 1126 CGTGGTGTCTACGCGCGGCGCAACAACTTGGAGAGATTTGGAGAACTCATGGCAG 1185
 Db 436 LysTrpPheTyrSerGlyHisLysProAsnLeuGluLysLeuAsnAlaTrpIle 455
 QY 1186 TCGATAGTGGGCGCTGTATGTTACCAATACCATATCTTCGAGTACAGATTCGTTACA 1245
 Db 456 SerIleSerGlyProValMetLeuTrpHisValPhePheArgValThrAspSerIleThr 475
 QY 1246 AAGAGAGACCGTCGACAGTTTGTACAAATACCAATGATTTAGTTCGTTGGTTCATCCTTCGTT 1305
 Db 476 ArgGluThrValGluSerLeuPheLysTyrHisAspLeuIleArgTyrSerThrIle 495
 QY 1306 CTGGGCTTGTGATGATTTGGNACTCGTGGAGAGGTGAGCAGAGGGATGTGCGG 1365
 Db 496 LeuArgLeuAlaAspLeuGlyThrSerLeuGluValSerArgGlyAspValPro 515
 QY 1366 AAATCACTTCAGTCTACATGAGTACTACAATGCTATCGAGCGCGGAGCGGAGCAC 1425
 Db 516 LysSerIleGlnCysTyrMetAsnAspAsnAlaSerGluGluAlaArgHis 535
 QY 1426 GTGAAATGGCTGATAGCGGAGGTGTGAAGAATGATTCGGAGAGGTGTGCAAGAT 1485
 Db 536 IleArgTrpIleAlaGluThrTrpLysLysIleAsnGluValTrpSerValAsp 555
 QY 1486 TCTCCATTCGCAAGATTTTATAGATGTGCAATGATTTAGAGAGGTGCGCAGTTG 1545
 Db 556 SerProPheCysLysAspPheIleAlaCysAlaAspMetGlyArgMetAlaGlnPhe 575
 QY 1546 ATGTACCAATATGAGATGGCCAGCGCACACACACCCCTATTATACATCAACATGACC 1605
 Db 576 MetTyrHisAsnGlyAspGlyHisGlyIleGlnAsnProGlnIleHisGlnMetThr 595
 QY 1606 AGAACCTTATTCGAG 1620
 Db 596 AspileLeuPheGlu 600

RESULT 9

Q9AXM7 PRELIMINARY; PRT; 604 AA.
 AC Q9AXM7;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Limonene synthase.
 OS *Perilla frutescens* var. *frutescens*.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Lamiales; Lamiaceae; Nepetoideae; Eisholtziaceae; Perilla.
 OX NCBI_TaxID=151328;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=5520;
 RA Takeya H., Ito M., Honda G.;
 RT "Limonene synthase from *Perilla frutescens*."
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF317695; AAK06663.1;
 DR HSSP; Q40577; SEAT.
 DR GO; GO:0016829; F:lyase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR005630; Terpene synth C.
 DR InterPro; IPR008949; Terpenoid synth.
 DR InterPro; IPR008930; Terp_cyc_toroid.
 DR InterPro; IPR001906; Terp_synth-like.

DR Pfam; PF01397; Terpene synth; 1.
 DR Pfam; PF01936; Terpene synth C; 1.
 SQ SEQUENCE 604 AA; 70528 MW; 7A8A8B3821774B59 CRC64;

Alignment Scores:

Prod. No.: 4,74e-149 Length: 604
 Score: 2002.00 Matches: 374
 Percent Similarity: 82.05% Conservative: 74
 Best Local Similarity: 69.50% Mismatches: 86
 Query Match: 68.84% Indels: 12
 DB: 10 Gaps: 3

US-09-938-956-6 (1-1632) x Q9AXM7 (1-604)

QY 4 AGAGCATCGGAAACTACAACTTCCTCGTGGGATGTCACCTCATTCATCGCTTCTC 63
 Db 61 ArgArgSerGlyAsnTyrSerProSerPheTrpAsnAlaAspTyrIleLeuSerLeuAsn 80
 QY 64 AGTGACTATTAAGGAGAC--AAACACGCTGATTTAGGGCTTCGAGTGGTCTACTTTGGTG 120
 Db 81 AsnHisTyrLysGluLysSerArgHisMetLysArgAlaGlyGluLeuIleValGlnVal 100
 QY 121 AGATGGAATCGGAGAAAGAAACGGATCAAATTCGACAACTTGTAGTTGATCATCATCTG 180
 Db 101 LysMetValMetGlyLysGluThrAspProValValGlnLeuGluLeuIleAspLeu 120
 QY 181 CAGAGATGGGGCTGCCGATCATTTCCAAAATAGTTCCAAAATAATCTTGTCTCTCTATA 240
 Db 121 HisLysLeuAlaLeuSerHisHisPheGluLysGluIleLysGluLeuPheAsnIle 140
 QY 241 TATCTCGACCATCACTATTACAAGAACCCCTTTCCAAAAGAAAGAAAGGATCTTACTCC 300
 Db 141 SerIleTyrAspHisLysIle-----MetValGluArgAspLeuTyrSer 155
 QY 301 ACATCTCTTGCAATTAGGCTCTCAGAGAAACATGTTTCAAGTCGCACAGAGATATTC 360
 Db 156 ThrAlaLeuAlaPheArgLeuLeuArgGlnTyrGlyPheLysValProGlnGluValPhe 175
 QY 361 GATAGTTTCAGAACGAGGAGGTGAGTTCAAGAAAGCCCTTAGCGACACACCGAGAGA 420
 Db 176 AspCysPheLysAsnAspAngLysGluPheLysArgSerLeuSerAspThrLysGly 195
 QY 421 TTCTTGCACATGTATGAAGCTTCCTTCTGTGTGAGGAGGCGGAAACCCAGCTCCAGTCA 480
 Db 196 LeuLeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluMetThrLeuGluLeu 215
 QY 481 GCGAGGATTCGCCACCAATTTTGGAGGAAAAGTGAACGAGGGTGGTGTGATCGC 540
 Db 216 AlaArgGluPheAlaThrIlePheLeuGlnGluLysLeuAsnAspLysThrIleAsp 235
 QY 541 GAC-----CTTTTAAACAAGATCGCATATTCTTTGACATCCCTCTT 582
 Db 236 AspAspAlaAspThrAsnLeuIleSerCysValArgHisSerLeuAspIleProIle 255
 QY 583 CATGAGGATTAAGGCCAATCCACCTGTGTGATCGATCGATGATATAGGAGAGGCC 642
 Db 256 HisTrpArgIleGlnArgProAsnAlaSerTrpTrpIleAspAlaTyrLysArgSer 275
 QY 643 GACATCAATCCAGTACTGTGTGGAGCTTGCATCTACTCTAAATATTGTTCAAGCAAA 702
 Db 276 HisMetAsnProLeuValLeuGluLeuAlaLysLeuAspLeuAsnIlePheGlnAlaGln 295
 QY 703 TTTCAAGAGAGCTCAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTGTGAGAG 762
 Db 296 PheGlnGlnGluLeuLysGlnAspLeuGlyTrpTrpLysAsnThrCysLeuAlaGluLys 315
 QY 763 CTGCCCTTCGCAAGGATAGACTGGTGGAAATGCTACTTTTGGAAATCTCGGATCATCGAG 822
 Db 316 LeuProPheThrArgAspArgLeuValGluCysTyrPheTrpCysThrGlyIleIleGln 335
 QY 823 CAGCTCAGCATGCAAGTCGAGGATATGATGGGCAAGTCAACGCTCTGATACGGTG 882
 Db 336 ProLeuGlnHisGluAsnAlaArgValThrLeuAlaLysValAsnAlaLeuIleThr 355

DR HSSP; Q40577; 5EAU.
DR GO; GO:0016829; P:lyase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005630; Terpenoid synth C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp cyc toxioid.
DR InterPro; IPR001906; Terp synth-like.
DR Pfam; PF01397; Terpene synth; 1.
DR Pfam; PF03936; Terpene synth; C; 1.
SQ SEQUENCE 603 AA; 70264 MW; 653EC5C7F2520AC8 CRC64;

Alignment Scores:
Pred. No.: 8,93e-149 Length: 603
Score: 1998.50 Matches: 374
Percent Similarity: 81.28% Conservative: 69
Best Local Similarity: 88.62% Mismatches: 91
Query Match: 68.72% Indels: 11
DB: 10 Gaps: 2

US-09-938-956-6 (1-1632) x Q9FV72 (1-603)

QY	4	AGACAGTCGGAACTACAAACCCCTCTCGTTGGATGTC	CAACTTCATCCAAATCGCTCTCTC	63
DB	61	ARGARGSRGlyAsnTyrrSerProSerPheTrpAsnAlaAsp	TrpLeuLeuLeuLeuAsn	80
QY	64	AGTGACTATAAGGAGGACAAACACGCTGATTAGGCGT	TCTGAGCTGGTCACTTTGGTGAAG	123
DB	81	SerHisTyrrLysAspLysSerHisMetLysArgAlaGly	GluLeuIleValGlnValLys	100
QY	124	ATGGAACTGGGAAGAAGAACCGATCAAAATTCGACA	ACTTCGAGTTGATGATCGATCGATTCGCAG	183
DB	101	MetValMetGlyLysGluThrAspProValGlnLeuGlu	LeuIleAspAspLeuGln	120
QY	184	AGGATGGGGCTGCTCCGATCAATTCAAAATAGATT	CAAGAAATCTTGTCCTCTATATAT	243
DB	121	LysLeuAlaLeuSerHisHisValGluLysGluIleLys	GluIleLeuPheLysIleSer	140
QY	244	CTCGACCATCACTATTACAGAACCTTTTCCAAAAG	AAAGAAAGGATCTCTACTCCACA	303
DB	141	ThrTyrrAspHisLysIle-	-----MetValGluArgAspLeuTyrrSerThr	155
QY	304	TCTCTTGATTTAGGCTCCTCAGAGAAACATGGTTT	CAAGTCCGACAAAGAGGTATTCGAT	363
DB	156	AlaLeuAlaPheArgLeuLeuArgGlnTyrrGlyPheLys	ValProGlnGluValPheAsp	175
QY	364	AGTTTCAGAACGAGGAGGTGAGTTCAAGAAAGCCTT	AGCAGCAGCACACCAGAGGATTG	423
DB	176	CysPheLysAsnAspAsnGlyGluPheLysArgSerLeu	SerSerAspThrLysGlyLeu	195
QY	424	TTGCAACTGTATGAAGCTTCCTTCTGTTCAGCGAAG	CGCAAAACACGCTCGAGTCAGCG	483
DB	196	LeuGlnLeuTyrrGluAlaSerPheLeuLeuThrGluGly	GluMetThrLeuGluLeuAla	215
QY	484	AGGGAATTCGCCACCAAAATTTTTCGAGGAAAAGT	GAACAGCGGTGTGTGTATGCGCAC	543
DB	216	ArgGluPheAlaThrLysSerLeuGlnGluLysLeuAsn	GluLysThrIleAspAspAsp	235
QY	544	-----CTTTTAAACAAGATCCATATCTTTTGACAT	CCCTCTTCAT	585
DB	236	AspAspAlaAspThrAsnLeuIleSerCysValArgHis	SerLeuAspIleProIleHis	255
QY	586	TGGAGATTAAAGGCCAAATGCACCTGTGTGGATCGAAT	TGGTATAGGAAGAGCGCCGAC	645
DB	256	TrpArgIleGlnArgProAsnAlaSerTrpTrpIleAsp	AlaTyrrLysArgArgSerHis	275
QY	646	ATGAATCCAGTAGTGTTCGAGCTTGCATATCTGCAT	TAAATATTGTTCAAGCAACAATT	705
DB	276	MetAsnProLeuValLeuGluLeuAlaLysLeuAspLeu	AsnIlePheGlnAlaGlnPhe	295
QY	706	CAAGAAGAGCTCAAGAATCCCTTCAGTGTGTGGAGAAAT	ACTCGGTTTGTGTGAGAAGCTGT	765
DB	296	GlnGlnGluLeuLysGlnAspLeuGlyTrpTrpLysAsn	ThrCysLeuAlaGluLysLeu	315

766	CCTTTCGACGAAGGATAGACTGGTGGAATGCTACTTTTGGAAATACTGGGATCATCGAGCCA	825
Dy		
316	ProPheValArgAspArgLeuValGluCysTyPheTrpCysThrGlyIleIleGlnPro	335
Dy		
826	CGTCAGCATGCCAAGTCGCAAGGATAATGATCGGCACAAAGTCAAAGCTCTGATTACGGTGATC	885
Dy		
336	LeuGLHisGluAsnAlaArgValThrLeuAlaLysValAsnAlaLeulleThrThrLeu	355
Dy		
886	GATGATATTATCATGTCTTATGCACTTATAGAAGAACTCGAACAAATCATCTACGACCTCAT	945
Dy		
356	AspAspiletyrAspValTyrglyThrLeuGluGluLeuLeuPheThrGluAlaIle	375
Dy		
946	CGAAGATGGGATATAAACTCAATCGACCAACTCCCGATTACATGCAACTGTGCTTCTTCT	1005
Dy		
376	ArgArgTrpAspValSerSerIleAspHisLeuProAsnIlyrMetGlnIleuCysPheLeu	395
Dy		
1006	GCACCTCAACAACCTTCGTCGATGATACATCGTCAGATGTTATGAAGAGAGAAGCGGTCAAC	1065
Dy		
396	AlaLeuAsnAsnPheValAspAspThrAlaTyAspValMetIlysglulysAspIleAsn	415
Dy		
1066	GTTATACCCTACCTGGCGCAATCGTGGTGTGATTGGCGGATAAGTATATGTTAGAGGCA	1125
Dy		
416	IleileProtyrLeuArgIlySerTrpIleuasPheUalagluThryrLeuValGluAla	435
Dy		
1126	CGGTGGTTCACGGCGGGCACAAACCAAGTTTGGAGAGTATTTGGAGAACTCATGCGCAG	1185
Dy		
436	LysTrpPheTyrseryGlyHisLysProAsnLeuGluGluTyrlleuasAsnAlaTrpIle	455
Dy		
1186	TGCATAAGTGGCGCCTGTATGTTAAACCATATCTTCGAGTAACAGAGTATGTTTCACA	1245
Dy		
456	SerIleSerglyProValMetLeuCysHisValPhePheArgValThrAspSerIleThr	475
Dy		
1246	AAGGAGACCGTCACAGCTTTGTACAAATACACAGATTAGTTCGTTGGTCATCCTTCGTT	1305
Dy		
476	ArgGluThrValGluSerLeuPheLysTyHisaspLeulleargtyrSerSerThrIle	495
Dy		
1306	CTCGGCTTCTGTANGATTGGGAACCTCGGTGAAGAGGTGACGAGGGGATGTCCCG	1365
Dy		
496	LeuArgLeuAlaAspAspLeuGlyThrSerLeuGluGluValSerArgIlyAspValPro	515
Dy		
1366	AAATCACTTCAGTGCTCATGAGTGACTACAATGCATCGAGGGCGGAGCGCGGAAGCAC	1425
Dy		
516	LysSerIleGlnCysTyrMetAsnAspAsnAsnAlaserGluGluAlaArgArgHis	535
Dy		
1426	GTGAAATGGCTGATAGCGGAGGTGTGAAGAAGATGAATCGGAGAGGGGTGTGGAAGGAT	1485
Dy		
536	IleArgTrpLeuilleAlaGluThrTrpIlysIlyleAsnGluGluValTrpSerValAsp	555
Dy		
1486	TCTCCATTGCGCAAGAATTTTAGGATGTGCACTTGTATTAGAAAGGATGGCGCAGTTG	1545
Dy		
556	SerProPheCysIlyAspPheIleAlaCysAlaalaspMetGlyArgMetAlaGlnPhe	575
Dy		
1546	ATGTACCATATGAGATGGCGCGGCACACCAACCCCTATTATACATCAACAAATGACC	1605
Dy		
576	MetTyrlHisenGlyAspGlyHisglylleGlnAsnProGlnIleHisGlnGlnMetThr	595
Dy		
1606	AGAACCTTATTCGAG	1620
Dy		
596	AspileLeuPheGlu	600
Dy		

[1]
SEQUENCE FROM N.A.
TISSUB=Glandular trichomes;
Hoelscher D.J., Wildung M.R., Williams D.C., Croteau R.;
"A cDNA clone for a 3-carene synthase from Salvia stenophylla.";
Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
EMBL; AF527416; AAM89254.1; -
GO; GO:0016829; F:llyase activity; IEA;
GO; GO:0008152; P:metabolism; IEA;
InterPro; IPR005630; Terpene_synth_C.
InterPro; IPR008949; Terpenoid_synth.
InterPro; IPR008930; Terp_cyc_toroid.
InterPro; IPR001906; Terp_synth-like.
Pfam; PF01397; Terpene_synth_1.
Pfam; PF03936; Terpene_synth_C; 1.
SQ SEQUENCE 597 AA; 69714 MW; 4B562A8967B0F2F4 CRC64;
Alignment Scores:
Pred. No.: 5,46e-148 Length: 597
Score: 1988.50 Matches: 367
Percent Similarity: 82.14% Conservative: 79
Best Local Similarity: 67.59% Mismatches: 94
Query Match: 68.38% Indels: 3
DB: 10 Gaps: 3
US-09-938-956-6 (1-1632) x Q8L5J7 (1-597)
QY 4 AGAGATCCGGAAACTACAAACCTTCGTGGGATGCACCTCATCCTCAATCGCTTCTC 63
Db 51 ArgArgSerGlyAsnTySerProSerLysTrpAspValAspTyrlleGlnSerLeuHis 70
QY 64 AGTGACTATAAGAGAGACAACACGTCATTAGGGCTTCGTAGCTGGTGCACCTTTGGTGAAG 123
Db 71 SerAspTyrllysGlulGuArGHISThrArgArgAlaSerGlulLeuIleMetGluVallys 90
QY 124 ATGGAACTGGAGAAAGAAAACGGATCAAATTCGACACTTGAGTTGATGACTGCTCGAG 183
Db 91 LysLeuLeuGluLysGluProAsnProThrArgGlnLeuGluLeuIleAspAspLeuGln 110
QY 184 AGGATGGGCTGCCGATCATTTCCAAATAGAGTTCAAAGAAATCTTGCTCTATATAT 243
Db 111 LysLeuGlyLeuSerAspHIsPheAsnAsnGluPheLysGluIleLeuAsnSerValtyr 130
QY 244 CTCGACCATCATATTACAGAACCCTTTCCAAAGAAA---GNAAGGGATCTCTACTCC 300
Db 131 LeuAspAsnLysTyrtyrArgAsnGlyAlaMetLysGluValGluArgAspLeutySer 150
QY 301 ACATCTCTTGATTTAGCTCTCTCAGAGAACATGGTTTTCAAGTCGCACAGAGGTATTC 360
Db 151 ThrAlaLeuAlaPheArgLeuLeuArgGlnHISglyPheGlnValAlaGlnAspValLeu 170
QY 361 GATAGTTTCAAGACGAGGAGGGTGAAGTTCAAAGAAACCTTAGCGACGACACACAGAGGA 420
Db 171 GluCysPheLysAsnThrLysGlyGluPheGluProSerLeuSerAspAspThrArgGly 190
QY 421 TTGTTGCAACTGTATGAAGCTTCCTTCTGTGTGAGGAAGCGCAACCCAGCTCGAGTCA 480
Db 191 LeuLeuGlnLeuTyrtyrGluAlaSerPheLeuLeuThrGluGlyGluAsnThrLeuGluLeu 210
QY 481 GCGAGGGAAATTCGCCACCAAAATTTTGGAGAAAAAGTGAACGAGGGGTGGTGTGATGGC 540
Db 211 AlaArgAspPheThrThrLysIleLeuGluGluLysLeuArgAsnAspGluIleAspAsp 230
QY 541 ---GACCTTTTAAACAAGAAATCGCATATTTCTTGACATCCCTCTTCATTTGGAGATTAAA 597
Db 231 IleAsnLeuValThrTrpIleArgHisSerLeuGluIleProIleHisTrpArgIleAsp 250
QY 598 AGGCCAAATGCACCTGTGTGATGATGAATGGTATAGGAAGAGCCGCGCATCAATCCAGTA 657
Db 251 ArgValAsnThrSerValTrpIleAspValTyrllysArgArgProAspMetAsnProIle 270
QY 658 GTGTTGGAGCTTGGCATACTCGACTTAATATTTGTTCAAGCACAAATTTCAAGAAGCTC 717


```

Db      271 ValLeuGluLeuAlaValLeuAspSerAsnIleValGlnAlaGlnTyrGlnGluGluLeu 290
QY      718 AAAGAATCTTCAGTGTGTGAGAAATACTGGTGTGTGAGAGCTGCCCTTCGCAAGG 777
Db      291 LysLeuAspLeuGlnTyrTrpArgAsnThrCysLeuAlaGluLysLeuProPheAlaArg 310
QY      778 GATGACTGGTGGGAATGCTACTTTGGGAATACTGGGATCATCGACCCACGTCACATGCA 837
Db      311 AspArgLeuValGluSerTyrPheTrpGlyValGlyValValGlnProArgGlnHisGly 330
QY      838 AGTCAAGATAATGATGGCAAGTCAACGCTCTGATTACGGTGATCGATGATATTAT 897
Db      331 IleAlaArgMetAlaValAspArgSerIleAlaLeuIleThrValIleAspAspValTyr 350
QY      898 GATGCTATGCACTTGAAGAAGCTCGAACAAATCTCACTGACCTCATTTCGAAGATGGAT 957
Db      351 AspValTyrGlyThrLeuGluGluLeuGlnPheThrGluAlaIleArgArgTrpAsp 370
QY      958 ATAACTCAATCGACCAACTCCCGATTACATGCAACTGTCTTCTTGTGCACTCAACAAC 1017
Db      371 IleSerSerIleAspGlnLeuProSerTyrMetGlnLeuCysPheLeuAlaLeuAspAsn 390
QY      1018 TCTGTCGATACATCTGATGATTATGAGGAGAAAGCGTCAACGTTATACCCCTAC 1077
Db      391 PheIleAsnAspIleAlaTyrAspValLeuLysGluGlnGlyPheAsnIleIleProTyr 410
QY      1078 CTGCGCAATCTGGGTGATTCTGGCGATGAAGTATATGAGGACGCGTGTCTTCTAC 1137
Db      411 LeuArgLysSerTrpThrAspMetIleGluGlyPheLeuLeuGluAlaLysTrpTyrHis 430
QY      1138 GCGGGGCAACAACCAAGTTTGAAGATTTTGGAGAACTCATGGCACTCATGGCAGTCGATAGTGG 1197
Db      431 AsnGlyHisLysProLysLeuGluGlnTyrLeuGluAsnGlyTrpArgSerIleGlySer 450
QY      1198 CCTGTATGTTAAACGCACATATCTTCGAGTAACAGATTTCGTTCAACAAGGACACGCTC 1257
Db      451 ThrValValLeuThrHisAlaPhePheGlyValThrHisSerLeuThrLysGluAsnIle 470
QY      1258 GACAGTTTGTCAAAATACACAGATTATGCTTGGTTCATCTTCGTTCTCGCGCTGCT 1317
Db      471 AspGlnPhePheGlyTyrHisGluIleValArgLeuSerSerMetLeuLeuArgLeuAla 490
QY      1318 GATGATTTGGAACTCGTGGAGAGGTGACGAGGAGGTGCGGAGGATGCGGAATCACTTCAG 1377
Db      491 AspAspLeuGlyThrSerThrAspGluValSerArgGlyAspValProLysAlaIleGln 510
QY      1378 TCTCATCATGAGTAC---TACAATGCATCGGAGCGGAGCGGCGGAGACGTCGAAATGG 1434
Db      511 CysTyrMetAsnAspAsnIleGlyAlaSerGluAlaGluAlaArgGluHisValLysTrp 530
QY      1435 CTGATAGCGGAGGTGTGAGNAGATGATGCGGAGAGGTTGTCGAGGATTCCTCCATTC 1494
Db      531 CysIleTrpGluThrTrpLysLysMetAsnLysValArgValAlaArgAspThrProPhe 550
QY      1495 GCGAAAGATTTTATAGGATGTGCACTTATAGGAAGGATGGCGCAGTTCATGTATACCAT 1554
Db      551 SerGlnAspPheIleValCysAlaMetGlyMetGlyArgMetGlyGlnTyrMetTyrHis 570
QY      1555 ANTGAGATGGCAGCGGACACACACCTTATATATACATCAACAATGACACAGACCTTA 1614
Db      571 TyrGlyAspGlyHisGlyIleGlnHisSerIleIleHisGlnMetSerThrCysLeu 590
QY      1615 TTCGAGCCC 1623
Db      591 PheHisPro 593

```

RESULT 12

Q9LKM5

ID Q9LKM5 PRELIMINARY; PRT; 608 AA.

AC Q9LKM5;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

```

DE      Terpene synthase.
OS      Perilla frutescens.
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC      Lamiales; Lamiaceae; Nepetoideae; Elsholtzieae; Perilla.
OX      NCBI_TaxID=48386;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=5526;
RT      Yagura T., Ito M., Honda G.;
RL      "Unknown terpene synthase from Perilla frutescens.";
DR      EMBL; AF271259; AAF76186.1; -
DR      HSSP; Q40577; SEAU.
DR      GO; GO:0016829; F:lyase activity; IEA.
DR      GO; GO:0008152; P:metabolism; IEA.
DR      InterPro; IPR005630; Terpene synth C.
DR      InterPro; IPR008949; Terpenoid synth.
DR      InterPro; IPR008930; Terp_cyc_Teroid.
DR      InterPro; IPR001906; Terp_synth-like.
DR      Pfam; PF01397; Terpene synth; 1.
DR      Pfam; PF03936; Terpene synth C; 1.
SQ      SEQUENCE 608 AA; 70836 MW; 6A4D49AD55276010 CRC64;

```

Alignment Scores:

```

Pred. No.:      5,72e-144      Length:      608
Score:          1937.50      Matches:      368
Percent Similarity: 81.42%      Conservative: 79
Best Local Similarity: 67.03%      Mismatches: 87
Query Match:      66.63%      Indels:      15
DB:              10      Gaps:      5

```

US-09-938-956-6 (1-1632) x Q9LKM5 (1-608)

```

QY      4 AGACGATCGGAACTACAAACCTCTCGTGGGATGTCAACTTCATCAATCGTTCTC 63
Db      58 ArgArgSerGlyAsnTyrSerProSerPheTrpAsnThrAspTyrIleLeuSerLeuAsn 77
QY      64 AGTCACTAATAGGAGAGCAACACGCTGATT---AGCGCTCTCGAGCTGCTCATTGGTG 120
Db      78 CysAspTyrGluAspGluArgMetArgGlyAlaAlaGlyGluLeuValGluGlnVal 97
QY      121 ACATCGAACTCGAGAAAGAAACGGATCAATTCGACAACTTGAGTTGATGATGATGCTTG 180
Db      98 LysMetLeuMetGluLysGluThrAspProIleValGlnLeuGluLeuLeuValLeu 117
QY      181 CAGAGGATGGGCTGTCGATCATTTCCAAATAGTTCAAAGAAATCTTG----- 231
Db      118 GlnLysLeuAlaLeuSerHisHisPheGluLysGluPheGluGlyIleLeuPheAsnIle 137
QY      232 TCCTCTATATATCTCGACCATCACTATTACAAGAACCTTTTCCAAAGAAAGAAAGGAT 291
Db      138 SerThrIleTyrAspAsp-----LysAsn-----ArgGluArgAsp 149
QY      292 CTCTACTCCACATCTCTTCATTTAGCTCCTCAGAGAACATCGTTTTCAGTCGACAA 351
Db      150 LeuTyrSerThrThrLeuAlaPheArgLeuLeuArgGlnHisGlyTyrGlnValProGln 169
QY      352 GAGGTATTCATAGTTTTCAGACGAGGAGGAGTTCAAAGAAAGCCTTACGACGAC 411
Db      170 GluLeuPheGluCysPheLysAsnAspLysGlyGluPheLysGluSerLeuSerAsnAsp 189
QY      412 ACCAGAGGATTGTTCGAACCTGTATGAAGCTTCCTTTCTGTTCGCGGAGCGGAAACCA 471
Db      190 ThrLysGlyLeuLeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThr 209
QY      472 CTGAGTCAGCGAGGAAATTCGCCCAATTTTTCGAGGAAAGAAAGTGAAC----- 522
Db      210 LeuGluLeuAlaArgGluPheAlaThrLysPheLeuGlnGlnLysGluLysLeuAsnIle 239
QY      523 GAGGCTGTGCTGATGCGGACCTTTTAAACAAGAAATCGCATATTCTTTGGACATCCCTCT 582
Db      230 AspAspAspAspThrAsnLeuIleSerCysValArgHisSerLeuAspMetProIle 249

```



```
Db 191 LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu 210
Qy 481 GCGAGGAATTCGCCACCAATATTTTGGAGGAAAGTGAACGAGGGTGGTGTGATGGC 540
Db 211 AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu-----ValAspLys 227
Qy 541 GAC-----CTTTAAACAAGATCGCATATTCCTTTGGACATCCCTCTTCATTGAGGATT 594
Db 228 AspLeuLeuLeuSerSerIleGluAlaLeuGluLeuProThrHisLysTrpArgVal 247
Qy 595 AAAAGGCCAAATGCACCTGTGGATCGAATGGTATAGGAAGAGCCCGACATCAATCCA 654
Db 248 GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgProAspMetAsnPro 267
Qy 655 GTAGTGTGGAGCTTGCCTACCTGACCTTAATATTTGTTCAAGCACATTTTCAAGAAGAG 714
Db 268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlnGlu 287
Qy 715 CTCGAAGAATCCTTCAGTGGTGGAGAAATACTGGTGTGTGTGAGAGCTCCCTTCGCA 774
Db 288 LeuLysGluAlaSerArgTrpTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307
Qy 775 AGGATAGACTGGTGGATGCTACTTTTGGAAATCTGGATCATCGAGCCACGTCAGCAT 834
Db 308 ArgAspArgIleValGluCysTyrTyrTrpThrThrGlyValValGluArgArgGluHis 327
Qy 835 GCAAGTCACAGATATGATGCGCAAGTCAACCTCTGATACCGTATCGATGATGATATT 894
Db 328 GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrIleAspAspVal 347
Qy 895 TATGATGCTATGCGACCTTGAAGAATCGAACTCAATTTCACTGACCTCATTCGAAGATGG 954
Db 348 PheAspIleTyrGlyThrLeuGluGluLeuGlnLeuPheThrThrAlaIleGlnArgTrp 367
Qy 955 GATATAACTCAATCGACCACTCCGATACATGCACTGTGTGTTCTTGGCTACTCAAC 1014
Db 368 AspIleGluSerMetLysGlnLeuProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387
Qy 1015 AACTTCGTCGATGATACATCGTACCATCTGTATGAAGGAGAAAGGCGTCAACGTTATACC 1074
Db 388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407
Qy 1075 TACTTCGGCAATTCGGTGGTATTTGGCGGATAAGTATATGTTAGAGCGACGGTGGTTC 1134
Db 408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTrpTyr 427
Qy 1135 TACGCGGGCACAAACCAAGTTTGAAGAGATTTTGGAGACTCATGCGACCTCGCATAGT 1194
Db 428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447
Qy 1195 GGGCCCTGTATGTTAAACGCACATATTTCTCCAGTAAACAGATTCGTTCAAAAGAGACC 1254
Db 448 GlyIleProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluAsp 467
Qy 1255 GTCCGACAGTTTGTACAAATACACCATTTAGTTTGGTGTATCTCTGTTCTGGCGTT 1314
Db 468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487
Qy 1315 GCTGATGATTTCGGGAACCTCGGTGAAGAGGTGACGAGGGGATGTCGCAAAATCACTT 1374
Db 488 AlaAspAspMetGlyThrSerLeuAspGluValGluArgGlyAspValproLysSerVal 507
Qy 1375 CAGTGTCTACATGAGTGTACATATGTCATCGAGGGCGGCGGCGGAACGACGTGAATGG 1434
Db 508 GlnCysTyrMetAsnGlnLysAsnAlaSerGluGluAlaArgGluHisValArgSer 527
Qy 1435 CTGATACCGAGGTGTGGGAAGATGATGATGCGGAGAGGGTGTGCAAGGATTCCTCATTC 1494
Db 528 LeuIleAspGlnThrTrpLysMetMetAsnLysGluMetMetThr-----SerSerPhe 545
Qy 1495 GCGCAAGATTTTATGAGTGTGTCAGTTGATTTAGGAAGATCGCCAGTGTGATGATAC--- 1551
Db 546 SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln 565
```

1552 CATAATGGAGATGGGCGACGCGCACACACCTTATTATATCATCAACAATGACCAAGACC 1611
HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585
1612 TTATTTCGAGCCCTTT 1626
586 LeuPheAspArgTyr 590

RESULT 14
081193 PRELIMINARY; PRT; 590 AA.
AC 081193
DT 01-NOV-1998 (T-EMBLrel. 08, Created)
DT 01-NOV-1998 (T-EMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE (-)-sabinene synthase.
OS Salvia officinalis (Sage).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Lamiaceae; Nepetoideae; Menthaeae; Salvia.
OC NCBI_TaxID=38868;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=98279000; PubMed=9614092;
RX Wise M.L., Savage T.J., Katahira E., Croteau R.;
RT "Monoterpene synthases from common sage (Salvia officinalis). cDNA
RT isolation, characterization, and functional expression of (+)-sabinene
RT synthase, 1,8-cineole synthase, and (-)-bornyl diphosphate synthase.";
RL J. Biol. Chem. 273:14891-14899(1998).
DR EMBL; AF051901; AAC26018.1; -.
DR HSSP; Q40577; 5E4U.
DR GO; GO:0016829; F-lyase activity; IEA.
DR GO; GO:0008152; P-metabolism; IEA.
DR InterPro; IPR005630; Terpene synth C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp_cyc_forold.
DR InterPro; IPR001906; Terp_synth-like.
DR Pfam; PF01397; Terpene synth, 1.
DR Pfam; PF03936; Terpene synth C; 1.
SQ SEQUENCE 590 AA; 65942 MW; 3E9CDFE2A2F0307D CRC64;

Alignment Scores:
Pred. No.: 1,01e-113 Length: 590
Score: 1553.50 Matches: 291
Percent Similarity: 72.48% Conservative: 104
Best Local Similarity: 53.39% Mismatches: 141
Query Match: 53.42% Indels: 9
DB: 3

US-09-938-956-6 (1-1632) x 081193 (1-590)

Qy 1 ATGACAGCATCCGAAACTACACCTTCTCGTTGGATGTCAACTTCATCCATCGCTT 60
Db 51 IleArgArgSerGlyAspTyrGlnProSerLeuTipAspPheAsnTyrIleGlnSerLeu 70
Qy 61 CTGATGATCTATAGGAGGACAAACGATGATGAGGCTTCTGAGTGTGCTGCTTGGTG 120
Db 71 AsnThrProTyrLysGluGlnArgHisPheAsnArgGlnAlaGluLeuMetGlnVal 90
Qy 121 AAGATGCACTCGAGAAAGAAACGATCAATTCGACAACTTGATTCGATGATGCTTGG 180
Db 91 ArgMetLeuLeuLysValLysMetGluAlaIleGlnGlnLeuGluLeuIleAspLeu 110
Qy 181 CAGAGATGGGTGTCGATCATTTTCCAAATAGATTCAAAGAAATCTTGTCTCTATA 240
Db 111 GlnTyrLeuGlyLeuSerTyrPhePheGlnAspGluIleLysGlnIleLeuSerSerile 130
Qy 241 TATCTCGACCATCACTATTACAGAACCCCTTTTCCAAAGAAAGGATCTCTACTCC 300
Db 131 HisAsnGluProArgTyrPheHisAsn-----AsnAspLeuTyrPhe 144
Qy 301 ACATCTCTGATTTAGGCTCTCTCAGAGAACATGGTGTTCAGTCCGACAAAGGATTC 360

Db 123 LysAsnLeuGlyLeuSerTyrPhePheGluAspGlnIleLysGlnIleLeuThrPheIle 142
 QY 241 TATCTCGACCATCAC--TATTACAAAGACCCCTTTT-----CCAAAGAGAAAGAGGAT 291
 Db 143 TyrAsnGluHisLysCysPheHisSerAsnSerIleIleGluAlaGluIleLeuArgAsp 162
 QY 292 CTCTACTCCACATCTCTTTCATTTAGTTCCTCAGAGAACATGGTTTCAATCGACACAA 351
 Db 163 LeuTyrPheThrAlaLeuGlyPheArgLeuLeuArgGlnHisGlyPheGlnValSerGln 182
 QY 352 GAGGTATTCGATAGTTTCAAGACGAGGAGGGT--GAGTTCAAGAAAGCCTTAGCGAC 408
 Db 183 GluValPheAspCysPheLysAsnGluGluGlySerAspPheLysAlaArgLeuGlyAsp 202
 QY 409 GACACAGAGGATGTGTCAACTGTATGAAGTTCCTTTCTGTGTGCGAAGCGCGAACC 468
 Db 203 AspThrLysGlyLeuLeuGlnLeuTyrGluAlaSerPheLeuLeuArgGluGlyGluAsp 222
 QY 469 ACCTCGAGTCAGCGAGGAATTCGCCACCAATTTTGGAGAAAGAGTGAACGAGGT 528
 Db 223 ThrLeuGluLeuAlaArgGlnTyrAlaThrLysPheLeuGlnLysLysValAspHisGlu 242
 QY 529 GGTGTT---GATCGCGACCTTTTAAACAAGAATCGCATATTTCTTTGACATCCTCTTCAT 585
 Db 243 LeuIleAspAsnAsnLeuLeuSerTrpIleLeuHisSerLeuGluIleProLeuHis 262
 QY 586 TCGAGGATTAAGGCGCAATGACCTGTGTGATCGAATGTGTATAGGAAAGGCGCCGAC 645
 Db 263 TrpArgIleGlnArgLeuGluAlaArgTrpPheLeuAspArgTyrAlaThrArgArgAsp 282
 QY 646 ATCAATTCAGTGTGTGGAGCTTGCCTCACTTAAATATTTTCAAGCACACAAATTT 705
 Db 283 MetAsnGlnIleLeuGluLeuAlaLysLeuAspPheAsnIleIleGlnAlaThrGln 302
 QY 706 CAAGAAGAGCTCAAGAATCCTTTAGTGTGTGAGAAATACTGGGTTTGTGAGAGCTG 765
 Db 303 GlnGluGluLeuLysAspLeuSerArgTrpTrpLysSerThrCysLeuAlaGluLysLeu 322
 QY 766 CCCTTCGACGAGGATGACCTGTGGATGCTACTTTTGGATATCTGGGATCATCGAGCCA 825
 Db 323 ProPheValArgAspArgLeuValGluSerTyrPheTrpAlaIleAlaLeuPheGluPro 342
 QY 826 CGTCAGCATCAAGTCAAGGATAATGATGGGCAAACTCAACGCTCTGATTACGGTGTATC 885
 Db 343 HisGlnTyrGlyTyrHisArgLysValAlaAlaLysIleIleThrLeuIleThrSerLeu 362
 QY 886 GATGATATTTATGATGCTATGGCACTTGAAGAAGTCAACAAATCTGACCTCATTT 945
 Db 363 AspAspValTyrAspIleTyrGlyThrLeuAspGluLeuGlnPheThrAspAlaIle 382
 QY 946 CGAAGATGGGATATAAATCAATCGACCACTTCCCGATTACATGCAACTGTGCTTTCTT 1005
 Db 383 GlnArgTrpAspThrGluSerIleSerArgLeuProTyrTyrMetGlnLeuPheTyrMet 402
 QY 1006 GCATCAACAACTTCGTGATGATATGATGATGATGATGATGATGATGATGATGATGATG 1065
 Db 403 ValLeuTyrAsnPheValSerGluLeuAlaTyrAspGlyLeuLysGlyLysGlyPheIle 422
 QY 1066 GTTATACCTACCTCGGCAATCGGTGTGATTTGGCGGATAAGTATATGATGATGATGATG 1125
 Db 423 ThrIleProTyrLeuGlnArgSerTrpAlaAspLeuValGluAlaTyrLeuLysGluAla 442
 QY 1126 CGGTGGTTTACGCGGCGCAACCAAGTTTGGAAAGATTTTGGAGAACTCATCGGACG 1185
 Db 443 LysTrpPheTyrAsnGlyTyrValProSerMetGluGluTyrLeuAsnAlaTyrIle 462
 QY 1186 TCGATAGTGGCCCTGTATGTTAAGCAGCATATTTCTCCGAGTACAGATTCGTTTCACA 1245
 Db 463 SerIleGlyAlaThrProValIleSerGlnValPhePheThrLeuAlaThrSerIleAsp 482
 QY 1246 AAGGAGACCGTTCGACAGTTTGTACAAATACCAAGTTTGTGTTGTTGTTGTTGTTGTT 1305
 Db 483 LysProValIleAspSerLeuTyrGluTyrHisArgIleLeuArgLeuSerGlyMetLeu 502

QY 1306 CTGGGCTTGTGATGATTTTGGGAACCTCGTGGAGAGGTGACACAGAGGGATGTCCCG 1365
 Db 503 ValArgLeuProAspAspLeuGlyThrSerProPheGluMetLysArgGlyAspValPro 522
 QY 1366 AAATCACTTCAGTCTACTACATGAGTACTACATGCTATCGGAGGCGGAGCGGAGCAC 1425
 Db 523 LysAlaIleGlnLeuTyrMetLysGluArgAsnAlaThrGluIleGluAlaGlnGluHis 542
 QY 1426 GTGAATGGCTGATAGCGGAGGTCTGGAAGAAGATGAATCGGAGAGGGTGTCCGAAGAT 1485
 Db 543 ValArgPheLeuIleArgGluAlaTrpLysGluMetAsnThrValThrThrAlaAlaAsp 562
 QY 1486 TCTCCATTCGCGCAAGATTTTATAGGATGTCCAGTTCATTTAGGAAGATGCGGAGTTG 1545
 Db 563 CysProPheThrAspAspLeuValAlaAlaThrArgAsnLeuGlyArgAlaAlaGlnPhe 582
 QY 1546 ATGTACCATATGAGATGGGACGCGCACACACACCCCTATTATACATCAACAATGACC 1605
 Db 583 MetTyrLeuAspGlyAspGlyAsn-----HisSerGlnLeuHisGlnArgIleAla 599
 QY 1606 AGAACCTTATTCGAGCCCTTTGCA 1629
 Db 600 CysLeuLeuPheGluProTyrAla 607

Search completed: February 25, 2004, 05:49:12
 Job time : 120 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 25, 2004, 05:36:43 ; Search time 21.5 Seconds
(without alignments)
7904.976 Million cell updates/sec

Title: US-09-938-956-6

Perfect score: 2508

Sequence: 1 atgagacgacgcggaacta.....tattogagcccttgatga 1632

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 283362

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2.1/USPTO_epool_p/US09938956/runat_24022004_144700_3053/app_query.fasta_1.1799
-DB=SwissProt_42 -QFMT=fastan -SUFFIX=rsp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pcpt -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09938956@cgn_1_21@runat_24022004_144700_3053 -NCPUL=6 -ICPU=3
-NO MAP -LARGQUERY -NEG SCORE=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	805.5	27.7	548	1	SEAS_TOBAC
2	796	27.4	554	1	DCS2_GOSAR
3	781	26.9	554	1	DCS4_GOSAR
4	776	26.7	554	1	DCS1_GOSAR
5	776	26.7	554	1	DCS1_GOSHI
6	740	25.4	555	1	DCS3_GOSAR
7	714	24.6	628	1	TSDI_ABIGR
8	697	24.0	601	1	CASS_RICCO
9	681.5	23.4	637	1	TSDB_ABIGR
10	672.5	23.1	627	1	TSDB_ABIGR
11	638	21.9	862	1	TASY_ABIGR
12	635	21.8	862	1	TASY_TAXCH
13	634.5	21.8	862	1	TASY_TAXBA
14	138.5	4.8	1169	1	YK82 YEAST
15	130.5	4.5	2869	1	RBPI_PLAYB
16	127.5	4.4	943	1	SYL_METKA
17	125.5	4.3	1030	1	Y07E_MYCPN
18	124	4.3	1939	1	MYH6_MESAU

19	123.5	4.2	466	1	SYC_CLOPE	Q8xhq5 clostridium
20	123	4.2	1713	1	LMA3_HUMAN	Q16787 homo sapien
21	123	4.2	2672	1	GCN1_YEAST	P33892 saccharomyc
22	122.5	4.2	4540	1	DYHC_PART2	Q27171 paramecium
23	120.5	4.1	1938	1	MYH6_RAT	P02563 rattus norv
24	120.5	4.1	3680	1	DMD_CANFA	O97592 canis famil
25	118	4.1	1875	1	MLP1_YEAST	Q02455 saccharomyc
26	117.5	4.0	725	1	AGAL1_YEAST	P33223 saccharomyc
27	117	4.0	3685	1	DMD_HUMAN	P11532 homo sapien
28	115	4.0	1140	1	YM96_YEAST	Q04893 saccharomyc
29	115	4.0	1315	1	CHAO_DROME	P12024 drosophila
30	114.5	3.9	795	1	LON_MYCGE	P47481 mycoplasma
31	114.5	3.9	1938	1	MYH6_MOUSE	Q02566 mus musculu
32	114.5	3.9	2555	1	FAPY_HUMAN	O05057 h probable
33	114	3.9	847	1	SYA_HELPY	P56452 helicobacte
34	114	3.9	1575	1	IQG2_HUMAN	P13576 homo sapien
35	113	3.9	1022	1	TPS3_YEAST	P38426 saccharomyc
36	113	3.9	1587	1	SUR2_CAEEL	Q10569 caenorhabdi
37	112	3.9	1163	1	YPT4_CAEEL	P41882 caenorhabdi
38	112	3.9	8545	1	ANCI_CAEEL	Q9n4m4 caenorhabdi
39	111.5	3.8	571	1	YD55_METJA	O58750 methanococc
40	111.5	3.8	1935	1	MYH7_HUMAN	P12883 homo sapien
41	111.5	3.8	3744	1	TRAI_YEAST	P38811 saccharomyc
42	111.5	3.8	6885	1	SNE2_HUMAN	Q8wxh0 homo sapien
43	110.5	3.8	720	1	VP19_BPAPS	Q9t1s9 bacterioph
44	110.5	3.8	856	1	CNRB_MOUSE	P23440 mus musculu
45	110.5	3.8	999	1	OKRP_RAT	Q63617 rattus norv

ALIGNMENTS

RESULT 1

SEAS_TOBAC STANDARD; PRT; 548 AA.

AC Q40577; (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Aristolochene synthase (EC 4.2.3.9) (5-epi-aristolochene synthase)
(EAS)
DE Nicotiana tabacum (Common tobacco)
OC Rukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 56-73.
RC STRAIN=cv. NK326; PubMed=1438319;
RX MEDLINE=93066390; PubMed=1438319;
RA Facchini P.J., Chappell J.;
RT "Gene family for an elicitor-induced sesquiterpene cyclase in
to tobacco";
RL Proc. Natl. Acad. Sci. U.S.A. 89:11088-11092(1992).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS), AND REVISIONS.
RX MEDLINE=97442533; PubMed=9295271;
RA Starks C.M., Back K., Chappell J., Noel J.P.;
RT "Structural basis for cyclic terpene biosynthesis by tobacco 5-epi-
aristolochene synthase";
RL Science 277:1815-1820(1997).
CC -1- FUNCTION: Catalyzes the cyclization of trans,trans-farnesyl
diphosphate (FPP) to the bicyclic intermediate 5-epi-
aristolochene, initial step in the conversion of FPP to the
sesquiterpenoid antifungal phytoalexin capsidiol.
CC -1- CATALYTIC ACTIVITY: Trans,trans-farnesyl diphosphate =
aristolochene + diphosphate.
CC -1- COFACTOR: Binds 3 magnesium ions per subunit.
CC -1- SUBUNIT: Monomer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- INDUCTION: By fungal elicitor.
CC -1- SIMILARITY: Belongs to the terpene synthase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; L04680; AAA19216.1; -
PIR; T03714; T03714.
PDB; SEAS; 15-OCT-97.
PDB; SEAT; 12-NOV-97.
PDB; SEAU; 08-APR-98.
DR InterPro; IPR008930; Terp_cyc_toroid.
DR InterPro; IPR001906; Terp_synth-like.
DR InterPro; IPR005630; Terpene_synth_C.
DR InterPro; IPR008949; Terpenoid_synth.
DR Pfam; PF01397; Terpene_synth_1.
DR Pfam; PF03936; Terpene_synth_C; 1.
KW Lyase; Magnesium; 3D-structure.
PROTON ACCEPTOR.

Y -> YYY (IN REF. 1).
K -> Q (IN REF. 1).
N -> S (IN REF. 1).
M -> R (IN REF. 1).
T -> I (IN REF. 1).
D -> E (IN REF. 1).
T -> M (IN REF. 1).

```

CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
DR DR EMBL; L04680; AAA19216.1; -
DR DR PIR; T03714; T03714.
DR DR PDB; SEAS; 15-OCT-97.
DR DR PDB; SEAT; 12-NOV-97.
DR DR PDB; SEAU; 08-APR-98.
DR DR InterPro; IPR008930; Terp_cyc_toroid.
DR DR InterPro; IPR001906; Terp_synth-like.
DR DR InterPro; IPR005630; Terpene_synth_C.
DR DR InterPro; IPR008949; Terpenoid_synth.
DR DR Pfam; PF01397; Terpene_synth_1.
DR DR Pfam; PF03936; Terpene_synth_C; 1.
KW KW Lyase; Magnesium; 3D-structure.
KW KW PROTON ACCEPTOR.
FT FT ACT_SITE 273 273
FT FT ACT_SITE 444 444
FT FT ACT_SITE 520 520
FT FT ACT_SITE 525 525
FT FT ACT_SITE 525 525
FT FT CONFLICT 42 42
FT FT CONFLICT 44 44
FT FT CONFLICT 55 55
FT FT CONFLICT 62 62
FT FT CONFLICT 73 73
FT FT CONFLICT 89 89
FT FT CONFLICT 388 388
FT FT CONFLICT 26 29
FT FT CONFLICT 36 57
FT FT CONFLICT 58 58
FT FT TURN 60 61
FT FT TURN 64 76
FT FT TURN 77 78
FT FT TURN 80 83
FT FT TURN 84 97
FT FT TURN 104 116
FT FT TURN 117 118
FT FT TURN 123 129
FT FT TURN 130 130
FT FT TURN 132 133
FT FT TURN 136 136
FT FT TURN 138 142
FT FT TURN 144 154
FT FT TURN 155 156
FT FT TURN 159 160
FT FT TURN 162 164
FT FT TURN 165 166
FT FT TURN 167 178
FT FT TURN 179 181
FT FT TURN 184 185
FT FT TURN 186 195
FT FT TURN 199 200
FT FT TURN 203 213
FT FT TURN 214 214
FT FT TURN 215 217
FT FT TURN 219 220
FT FT TURN 223 254
FT FT TURN 255 255
FT FT TURN 256 259
FT FT TURN 261 262
FT FT TURN 267 277
FT FT TURN 281 283
FT FT TURN 284 306
FT FT TURN 307 307
FT FT TURN 310 322
FT FT TURN 323 323
FT FT TURN 326 330
FT FT TURN 333 354
FT FT TURN 355 357
FT FT TURN 359 361
FT FT TURN 361 361
FT FT TURN 362 385
FT FT TURN 386 386
FT FT TURN 391 398
FT FT TURN 399 399
FT FT TURN 400 402
FT FT TURN 404 413
FT FT TURN 414 414
FT FT TURN 416 417
FT FT TURN 420 427
FT FT TURN 428 428
FT FT TURN 431 454
FT FT TURN 455 456
FT FT TURN 458 459
FT FT TURN 461 469
FT FT TURN 470 470
FT FT TURN 473 494
FT FT TURN 495 495
FT FT TURN 503 505
FT FT TURN 506 506
FT FT TURN 507 519
FT FT TURN 520 520
FT FT TURN 535 542
FT FT TURN 543 543
SQ SQ SEQUENCE 548 AA; 62973 MW; 9FE1C59CF1A68BF1 CRC64;
Alignment Scores:
Pred. No.: 4.35e-51 Length: 548
Score: 805.50 Matches: 178
Percent Similarity: 52.92% Conservative: 112
Best Local Similarity: 32.48% Mismatches: 207
Query Match: 27.70% Indels: 51
DB: 1 Gaps: 10
US-09-938-956-6 (1-1632) x SEAS_TOBAC (1-548)
QY 1 ATGAGACGATCGGAAACTACACCTCTCTCGTTGGATGTCAACTTTCATCAATCGCTT 60
Db 14 ValArgProValAlaAspPheSerProSerLeuTyrGlyAspGlnPheLeuSerPheSer 33
QY 61 CTGAGTCACTATAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGCTGCTGGTG 120
Db 34 IleAspAsnGlnValAlaGluLysTyrAla----- 43
QY 121 AAGATGGAACCTGGAGAAAGAAACGGATCAAATTCGACAA----- 159
Db 44 ---LysGluIleGluAlaLeuLysGluGlnThrArgAsnMetLeuLeuAlaThrGlyMet 62
QY 160 -----CTTGAGTTGATCGATGCTGTCGAGAGATGGGCTGTCGATCAT 204
Db 63 LysLeuAlaAspThrLeuAsnLeuIleAspThrIleGluArgLeuGlyIleSerTyrHis 82
QY 205 TTCCAAATGAGTTCAAAGAAATCTTGTCTCTATATATCTCGACCATCACTATTACAAG 264
Db 83 PheGluLysGluIleAspAspIleLeuAspGlnIleTyrAsnGlnAsn----- 98
QY 265 AACCTTTTCCAAAGAAAGAGGATCTCTACTCCACATCTCTGCACTTTAGGCTCCTC 324
Db 99 -----SerAsnCysAsnAspLeuCysThrSerAlaLeuGlnPheArgLeuLeu 114
QY 325 AGAGAAACATGGTTTTCAGTCGCAAGAGGATTCGATAGTTTCAAGAACCGAGGAGGT 384
Db 115 ArgGlnHisGlyPheAsnIleSerProGluIlePheSerLysPheGlnAspGlnAsnGly 134
QY 385 GAGTTCAAAGAACCTTAGCGACGACACAGGATTTGTCACACTTTCATGAACTTCC 444
Db 135 LysPheLysGluSerLeuAlaSerAspValLeuGlyLeuLeuAsnLeuTyrGluAlaSer 154
QY 445 TTTCCTGTTCACGAAGCGAAACACCGCTCGAGTCAGGAGGAAATTCGCCACCAATTT 504
Db 155 HisValArgThrHisAlaAspAspIleLeuGluAspAlaLeuAlaPheSerThrIleHis 174
QY 505 TTGAGGAAAAAAGTGAACGAGGCTGTTGATGGGACCTTTTAAACAGATCCATAT 564
Db 505 TTGAGGAAAAAAGTGAACGAGGCTGTTGATGGGACCTTTTAAACAGATCCATAT 564

```

175 LeuGluSerAlaAlaProHis-----LeuLysSerProLeuArgGluGluValThrHis 192
 565 TCITTTGACATCCCTCTTCATGGAGGATTAAGGCGCAAAATGCACCTGTGTGATC-- 621
 193 AlaLeuGluGlnCysLeuHisLysGlyValProArgValGluThrArgPhePheLeuSer 212
 622 GAATGGTATAGGAAGAGCGCCGACATGAATCCAGTAGTGTGGAGCTTGCCATACCTGAC 681
 213 SerIleTyrAspLysGluGlnSerLysAsnValLeuLeuArgPheAlaLysLeuAsp 232
 682 TTAATATTGTTCAAGCACAAATTTCAAGAGAGCTTCAAGAACTTCCTCAGGTGGTGGAGA 741
 233 PheAsnLeuLeuGlnMetLeuHisLysGlnGluLeuAlaGlnValSerArgTrpTrpLys 252
 742 AATAGCTGGTTTGTAGAGAGCTCCCTTCGAGGAGTAGACTGTGGTAATCCTACTTT 801
 253 AspLeuAspPheValThrLeuProTyrAlaArgAspArgValValGluCysTrpPhe 272
 802 TGGAACTAGGATCATCGAGCGACGTCAGCATCGAAGTCAAGGATAAATGATGGGCAA 861
 273 TrpAlaLeuGlyValTyrPheGluProGlnTyrSerGlnAlaArgValMetLeuValLys 292
 862 GTCAAGCTTCGATACGCTGATCGATGATATTTATGATGCTATGTCACCTTAGAGAA 921
 293 ThrIleSerMetIleSerIleValAspAspThrPheAspAlaTyrGlyThrValLysGlu 312
 922 CTCGAACAATTCATGACCTCATTCGAAGATGGATATAACTCAATCGACCACTTCCC 981
 313 LeuGluAlaTyrThrAspAlaIleGlnArgTrpAspIleAsnGluIleAspArgLeuPro 332
 982 GATTACATGCACTGCTGCTTCTGCACTCAACAACTTCGTCGATGATACATCGATCAT 1041
 333 AspTyrMetLysIleSerTyrLysAlaIleLeuAspLeuTyrLysAspTyrGluLysGlu 352
 1042 GTTATGAAGGAGAAGGCTCAAGCTTATACCTTACCTCGGCAATCGTGGTGTGATTTG 1101
 353 LeuSerSerAlaGlyArgSerHisIleValCysHisAlaIleGluArgMetLysGluVal 372
 1102 CGGATATAGTATATGTTAGAGCGAGCTGGTCTTACCGCGCGGCACAAACCAAGTTTGGAA 1161
 373 ValArgAsnTyrAsnValGluSerThrTrpPheLeuGluTyrTrpProProValSer 392
 1162 GAGTATTTGGAGAACTCATGCGAGTCGATAGTAAAGTGGCGCTGTATGTTAAAGCACAATTC 1221
 393 GluTyrLeuSerAsnAlaLeuAlaThr-----ThrThrTyrTyr 405
 1222 TTCGAGTAACAGATTCGTTCTC-----ACAAAGGAGACCGTCGACAGT 1263
 406 TyrLeuAlaThrThrSerTyrLeuGlyMetLysSerAlaThrGluGlnAspPheGluTrp 425
 1264 TTGTACAAATACCAAGTATGTTAGTTCGTTGTCATCTTCGTTCTCGGCTTGTGATGAT 1323
 426 LeuSerLysAsnProLysIleLeuGluAlaSerValIleIleCysArgValIleAspAsp 445
 1324 TTGGGAACCTCGTGGAGAGGTGAGCAGAGGGGATGTCGCGAAATCACTTCAGTGTGATC 1383
 446 ThrAlaThrTyrGluValGluLysSerArgGlyGlnIleAlaThrGlyIleGluCysCys 465
 1384 ATGAGTGACTACAAATGCAATCGAGCGGCGAGCGGAGCAGCGTCAATGCTGATAGCG 1443
 466 MetArgAspTyrGlyIleSerThrLysGluAlaMetAlaLysPheGlnAsnMetAlaGlu 485
 1444 GAGGTGTGGAAGAAGATGAATCGGAGAGGGTGTGCAAGGATTTCTCCATTCGGCAAGAT 1503
 486 ThrAlaTrpLysAspIleAsn---GluGlyLeuLeuArgProThrProValSerThrGlu 504
 1504 TTTATAGATGTCGATGATTTAGAGAGGATGGCGAGTGTGATGATC---CATAATGGA 1560
 505 PheLeuThrProIleLeuAsnLeuAlaArgIleValGluValThrTyrIleHisAsnLeu 524
 1561 GATGGCGAGCGCACAAACACACCT 1584
 525 AspGlyTyr-----ThrHisPro 530

RESULT 2
 DCS2 GOSAR
 ID DCS2 GOSAR STANDARD; PRT; 554 AA.
 AC Q29760;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE (+)-delta-cadinene synthase isozyme XC14 (EC 4.2.3.13) (D-cadinene synthase).
 OS Gossypium arboreum (Tree cotton).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosid II; Malvales; Malvaceae; Malvoideae; Gossypium.
 OX NCBI_TaxID=29729;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nanking;
 RX MEDLINE=96132653; PubMed=8554317;
 RA Chen X.-Y., Chen Y., Heinsteins P., Davison V.J.;
 RT "Cloning, expression, and characterization of (+)-delta-cadinene synthase: a catalyst for cotton phytoalexin biosynthesis.";
 RL Arch. Biochem. Biophys. 324:255-266(1995).
 CC -!- FUNCTION: Responsible for the cyclization of trans,trans-farnesyl diphosphate (FPP) to (+)-delta-cadinene.
 CC -!- CATALYTIC ACTIVITY: 2-trans,6-trans-farnesyl diphosphate = (+)-delta-cadinene + diphosphate.
 CC -!- PATHWAY: Phytoalexins gossypol and lacinilene C biosynthesis;
 CC -!- first (committed) step.
 CC -!- SIMILARITY: Belongs to the terpene synthase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC EMBL; U23205; AAA93065.1; --
 DR PIR; S68366; S68366.
 DR HSP; Q40577; SEAU.
 DR InterPro; IPR008930; Terp_cyc_toroid.
 DR InterPro; IPR001906; Terp_synth-like.
 DR InterPro; IPR005630; Terpene synth C.
 DR InterPro; IPR008949; Terpenoid synth.
 DR Pfam; PF01397; Terpene synth; 1.
 DR Pfam; PF03936; Terpene synth; 1.
 DR Lyase; Multigene family.
 DR ACT_SITE 451 451 BY SIMILARITY.
 DR ACT_SITE 527 527 BY SIMILARITY.
 DR ACT_SITE 531 531 BY SIMILARITY.
 DR ACT_SITE 531 531 BY SIMILARITY.
 SQ SEQUENCE 554 AA; 64158 MW; A88974665B0F6B2B CRC64;
 Alignment Scores:
 Pred. No.: 2,15e-50 Length: 554
 Score: 796.00 Matches: 185
 Percent Similarity: 52.07% Conservative: 104
 Best Local Similarity: 33.33% Mismatches: 232
 Query Match: 27.37% Indels: 34
 DB: 1 Gaps: 7
 US-09-938-956-6 (1-1632) x DCS2_GOSAR (1-554)
 QY 1 ATGAGAGCATCCGAAACTACACCTTCTCGTTGG----- 36
 Db 21 MetArgProLysAlaAspPheGlnProSerIleTrpGlyAspLeuPheLeuAsnCysPro 40
 QY 37 GATGTCACTTCATCCAACTCGCTTCTCAGTACTATAAGGAGGACAAACACGCTGATTAG 96
 Db 41 AspLysAsnIleAspAlaGluThrGluLysArgHisGlnLeuLysGluGluValArg 60
 QY 97 GCTTCTGAGCTGCTCACTTTGGTGAAGATGGAACCTGAGAAAGAAACGGAATCAATCGA 156


```
Db 61 -----lysMetIleValAlaProMetAlaAsnSerThrGln 72
QY 157 CAACCTGAGTGCATGACATTCCTACATCTCTTGTGATTTAGCTCTCTCAGAGAACATGGT 216
Db 73 LysLeuAlaPheIleAspSerValGlnArgLeuGlyValSerTyrHisPheThrLysGlu 92
QY 217 TTCAGAAGAATCTGTCCTATATATCTCGACCATCACTATTACAAGAACCTTTTCCA 276
Db 93 IleGluAspGluLeuGluAsnIleTyrHisAsnAsn----- 105
QY 277 AAGAAGAAAGGATCTCTACCTCAATCTCTTGTGATTTAGCTCTCTCAGAGAACATGGT 336
Db 106 AspAlaGluAsnAspLeuTyrThrThrSerLeuArgPheArgLeuLeuArgGluHisGly 125
QY 337 TTTCAGTCCGCAACAGAGTATTCTGATCTTCAAGAACGAGGAGGTGAGTTCAAAGAA 396
Db 126 PheAsnValSerCysAspValPheAsnLysPheLysAspGluGlnGlyAsnPheLysSer 145
QY 397 AGCTTATGACGACACACAGAGGATTTGCAACTGTATGAAGCTCTCTTGTGACG 456
Db 146 SerValThrSerAspValArgGlyLeuLeuGluLeuTyrGlnAlaSerTyrLeuArgVal 165
QY 457 GAAGGGAACACAGCTCGAGTCAGGAGGATTCGCCACCAATTTTGGAGGAAAA 516
Db 166 HisGlyGluAspIleLeuAspGluAlaIleSerPheThrThrAsnHisLeuSerLeuAla 185
QY 517 GTCAACGAGGCTGTGTGATGTCGACCTTTTAAACAAGAAATCGCATATTCTTGGACATC 576
Db 186 Val-----AlaSerLeuAspTyrProLeuSerGluGluValSerHisAlaLeuLysGln 203
QY 577 CCTCTTCATTTGAGGAGTAAAGCCAAATGACCTGTGTGATCGAATGGTATAGGAAG 636
Db 204 SerIleArgArgGlyLeuProArgValGluAlaArgHisTyrIleuSerValTyrGlnAsp 223
QY 637 AGGCCGACATGAATCCAGTAGTGTGGAGCTGTCATCTACCTCACTTAATATTTGTCAA 696
Db 224 IleGluSerHisLysValLeuGluPheAlaLysIleAspPheAsnMetValGln 243
QY 697 GCACAAATTTCAAGAAGAGCTCAAGATCCTTCAGTGTGGAGAAATACCTGGTTTGT 756
Db 244 LeuLeuHisArgLysGluLeuSerGluLeuSerArgTyrTrpLysLeuAspPheGln 263
QY 757 GAGNAGTGCCTTCGCAAGGATAGACTGTGTGATGCTACTTTTGGATATCTGGATC 816
Db 264 ArgLysLeuProTyrAlaArgPheValGluGlyTyrPheTrpIleSerGlyVal 283
QY 817 ATCCAGCCAGCTCAGCATGCAAGTCAAGGATAATGATGCGCAAGTCAAGCTCTGATT 876
Db 284 TyrPheGluProGlnTyrSerLeuGlyArgLysMetLeuThrLysValIleAlaMetAla 303
QY 877 ACGGTGATCGATGATATTTATGATGCTATGGACCTTAGAAGACTCGAACAATTCAT 936
Db 304 SerIleValAspAspThrTyrAspSerTyrAlaThrTyrGluGluLeuIleProTyrThr 323
QY 937 GACCTCATTCGAAGATGGATATAAATCAATCAACCAACTCCCGATTCATCGCAACTG 996
Db 324 LysAlaIleGluArgTrpAspIleLysCysIleAspGluLeuProGluTyrMetLysPro 343
QY 997 TGTCTTCTGCACTCAACATTCCTCGATGATACATCGTACGATGTTATGAGGAGGAAA 1056
Db 344 SerTyrLysAlaLeuLeuAspValTyrGluGluMetGluGlnLeuValAlaLysHisGly 363
QY 1057 GCGGTCAACGTTATACCTACTCGGCAATCGTGGGTGATTTGGCGGATATAGTATATG 1116
Db 364 ArgGlnTyrArgValGluTyrAlaLysAsnAlaMetIleArgLeuAlaGlnSerTyrLeu 383
QY 1117 GTAGAGCACGGTGTCTTACCGCGGGCACAACCAAGATTGGAGAGTATTGGAGAAC 1176
Db 384 ValGluAlaArgTrpThrLeuGlnAsnTyrLysProSerPheGluGluPheLysAlaAsn 403
QY 1177 TCATGGCAGTCATAGTGGGCGCTGTATGTTAAACGCACATATCTTCCGAGTAAACAGAT 1236
```

```
Db 404 AlaLeuProThrCysGlyTyrAlaMetLeuAlaIleThrSerPheValGlyMetGlyAsp 423
QY 1237 TCCTTCAAAAAGAGACCGTCGACAGCTGTGTACAAATACACAGATTAGTTCGTGGTCA 1296
Db 424 IleValThrProGluThrPheLysTrpAlaAlaAsnAspProLysIleIleGlnAlaSer 443
QY 1297 TCCTTCGCTTCGGCTTCTGCTGATGATTTGGGAACCTCGTGGGAAGAGTGCACAGAGG 1356
Db 444 ThrIleLysCysArgPheMetAspValAlaGluHisLysPheLysHisArgArgGlu 463
QY 1357 GATGTGCGGAAATCACTTCAGTCTCATGAGTACTACAAATCATCGGAGCGGAGGCG 1416
Db 464 AspAspCysSerAlaIleGluCysTyrMetGluGluTyrGlyValThrAlaGlnGluAla 483
QY 1417 -----CGGAAGCACGCGAATGCTGCTGATGCGGAGGTGCGGAGAGATGAT 1464
Db 484 TyrAspValPheAsnLysHisValGlu-----SerAlaTrpLysAspValAsn 499
QY 1465 GCGGAGAGGCTGTCGAGGATTCCTCATTCGCAAAAGATTTTATAGGATTCGAGTTGAT 1524
Db 500 LysGluPheLeu---LysProThrThrGluMetProThrGluValLeuAsnArgSerLeuAsn 518
QY 1525 TTAGGAAGATGCGCGAGTGTGATGATACCATTAATGAGAGTGGCAGCGCACACACCCCT 1584
Db 519 LeuAlaArgValMetAspValLeuTyrArgGluGlyAspGlyTyrThrTyrValGlyLys 538
QY 1585 ATTATACATCAACAAATGACCAAGACCTTATTCGAGCCCTTTCGA 1629
Db 539 AlaAlaLysGlyGlyIleThrSerLeuLeuIleGluProValAla 553

RESULT 3
DCS4 GOSAR
ID DCS4 GOSAR STANDARD; PRT; 554 AA.
AC O49853;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, last sequence update)
DT 15-MAR-2004 (Rel. 43, last annotation update)
DE (+)-delta-cadinene synthase isozyme C2 (EC 4.2.3.13) (D-cadinene
synthase).
GN CADI-C2.
OS Gossypium arboreum (Tree cotton).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicot; rosids;
OC eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
OX NCBI_TaxID=29729;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nanking;
RA Meng Y., Jia J., Liu C., Liang W., Zhou X., Heinstein P., Chen X.-Y.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Responsible for the cyclization of trans,trans-farnesyl
CC diphosphate (FPP) to (+)-delta-cadinene.
CC -!- CATALYTIC ACTIVITY: 2-trans,6-trans-farnesyl diphosphate = (+)-
CC delta-cadinene + diphosphate.
CC -!- PATHWAY: Phytoalexins biosynthesis;
CC first (committed) step.
CC -!- SIMILARITY: Belongs to the terpene synthase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y16432; CAA76223.1; .
CC HSP; O40577; SEAU.
CC InterPro; IPR008930; Terp_cyc_toroid.
CC InterPro; IPR001906; Terp_synth-like.
CC InterPro; IPR005630; Terpene_synth_C.
CC InterPro; IPR008949; Terpeneoid_synth.
CC Pfam; PF01397; Terpene_synth; 1.
```

DR Pfam; PF09336; Terpene synth_C; 1.
 KW Lyase; Multigene family.
 FT ACT_SITE 451 BY SIMILARITY.
 FT ACT_SITE 527 BY SIMILARITY.
 FT ACT_SITE 531 BY SIMILARITY.
 SQ SEQUENCE 554 AA; 64117 MW; 35DD66D3E838AAC CRC64;

Alignment Scores:

Pred. No.: 2,67e-49 Length: 554
 Score: 781.00 Matches: 184
 Percent Similarity: 51.33% Conservative: 105
 Best Local Similarity: 32.68% Mismatches: 224
 Query Match: 26.86% Indels: 50
 Gaps: 9

US-09-938-956-6 (1-1632) x DCS4_GOSAR (1-554)

QY 1 ATGAGACGATCGGAACTACAACTCTCTCGTTGG----- 36
 Db 21 lleArgProLysAlaAspPheGlnProSerileTrpGlyAspPheLeuAsnCyPro 40
 QY 37 GATGTCACCTCATCGCTCTCTAGTACATATAGGAGGACAAACACGCTGATTAGG 96
 Db 41 AspLysAsnileAspAlaGlyThrGluLysArgHisGlnGlnLeuLysGluValArg 60
 QY 97 GCTTCTGAGCTGCTCACTTTGGTGAAGATGGAACCTGGAGAAAGAAACGATCAAAATCGA 156
 Db 61 -----LysMetileValAlaProMetAlaAsnSerThrGln 72
 QY 157 CAACCTGATGATCGATCGACGATGAGAGGATGGGCTGCTCGATCATCTCCAAATGAG 216
 Db 73 LysLeuAlaPheileAspSerValGlnArgLeuGlyValSerThrHisPheThrLysGlu 92
 QY 217 TTCGAAGAAATCTGCTCTATATATCTCGACCATCACTATTACAAAGAACCTTTTCCA 276
 Db 93 lleGluAspGluLeuGluAsnileTrpHisAsnAsn----- 105
 QY 277 AAAGAGAAAGGATCTCTACTCCATCTCTTGCAATTTAGCTCTCTCAGAGAACATGGT 336
 Db 106 AspAlaGluAsnAspLeuTrpThrSerLeuArgPheArgLeuLeuArgGluHisGly 125
 QY 337 TTTCAGTCGCACAGAGTATTCAGTAGTTTCAAGACGAGGAGGAGTCAAAAGAA 396
 Db 126 TyrAsnValSerCysAspValPheAsnLysPheLysAspGluGlnGlyAsnPhelysSer 145
 QY 397 AGCTTAGCGACGACACAGAGATTTGTCAACTGTATGAAGCTTCTCTGTGTGACG 456
 Db 146 SerValThrSerAspValGlnGlyLeuLeuGluLeuTyTrpGlnAlaSerThrLysArgVal 165
 QY 457 GAAGGGAACCAACGCTCAGTCAGGAGGAATTCGCCACCAAAATTTTGGAGGAAAAA 516
 Db 166 HisGlyGluAspilleLeuAspGluAlaileSerPheThrThrAsnHisLeuSerLeuAla 185
 QY 517 GTGACGAGGGTGGTGTGATCGGACCTTTTAAACAAGATCGCATATTTCTTTGGACATC 576
 Db 186 ValSer-----SerLeuAspHisProLeuSerGluGluValSerHisAlaLeuLysGln 203
 QY 577 CTTCTTCATTGGAGGATTAAGCCAAATGCACTGTGTGATCGAATGGTATAGGAG 636
 Db 204 SerIleArgGlyLeuProArgValGluAlaArgHisIleThrLeuSerValTyTrpGlnAsp 223
 QY 637 AGCCCGCATGATCCAGTAGTGTGGAGCTTGCATCTCGATCTAAATATTTCTCAA 696
 Db 224 IleGluSerHisAsnLysAlaLeuGluPheAlaLysileAspPheAsnMetLeuGln 243
 QY 697 GCACAAATTCAGAGAGCTCAAGAAATCTCTCAGGTGGTGGAGAAATCTGGTTGTT 756
 Db 244 PheLeuHisArgLysGluLeuSerGluileCysArgTiptTplyAspLeuAspPheGln 263
 QY 757 GAGAAGTCCCTTCGACGGGATAGACTGGTGGAAATGCTACTTTGGATATCTGGATC 816
 Db 264 ArgLysLeuProTyTrpAlaArgAspValValGluGlyTyTrpPheTrpIleSerGlyVal 283

QY 817 ATCGAGCCACGTCAGCATGCAAGTCAAGGATAATGATGGCAAAAGTCAACGCTCTGATT 876
 Db 284 TyrPheGluProGlnTrpSerLeuGlyArgLysMetLeuThrLysValleAlaMetAla 303
 QY 877 ACGGTGATCGATGATATTTATGATGCTATGACACCTTAGAAGAACTCGAAACAATTCAT 336
 Db 304 SerIleValAspAspThrTyTrpAspSerTyAlaThrTyTrpGluGluLeuileProTyTrp 323
 QY 937 GACCTCATTCGAGATGGGATATAAATCAATCGCACTTCCGATATACATCAACTG 996
 Db 324 AsnAlaileGluArgTrpAspIleLysCysileAspGluLeuProGluTyTrpMetLysPro 343
 QY 997 TGCTTTCTTTCGACTCAACAACTCTGTCGATGATACATCGTACGATGTTATGAGGAGAAA 1056
 Db 344 SerTyTrpLysAlaLeu-----LeuAspValTyTrpLysGluMet 355
 QY 1057 GCGCTCAACGTT-----ATACCTTACCTCGGCAATCGTGG 1092
 Db 356 GluGlnLeuValAlaGluHisGlyArgGlnTyArgValGluTyTrpLysAsnAlaMet 375
 QY 1093 GTTGATTTGCGGATAAGTATATGAGGACGCTGCTTACCGCGGGGACAAACCA 1152
 Db 376 IleArgLeuAlaGlnSerTyTrpLeuValGluAlaArgTiptThrLeuGlnAsnTyTrpLysPro 395
 QY 1153 AGTTTGGAAAGATTTTGGAGAACTCATGCGAGTGCATAGTGGGCGCTGTATGTTAACG 1212
 Db 396 SerPheGluGluPheLysAlaAsnAlaLeuProThrCysGlyTyTrpAlaMetLeuAlaile 415
 QY 1213 CACATATTTCTCCGAGTAAACAGATTCGTTCAACAAGGAGACCGCTCGACAGTTTGTACAA 1272
 Db 416 ThrSerPheValGlyMetGlyAspIleValThrProGluThrPheLysTrpAlaAlaAsn 435
 QY 1273 TACCAGATTTAGTTCGTTGTCATCTCTGCTTCGCTGCTGCTGCTGATGATTTGGGAACC 1332
 Db 436 AspProLysileileGlnAlaSerThrIlelleCysArgPheMetAspValAlaGlu 455
 QY 1333 TCGTGGAGAGTGGAGGAGGAGGATGTCGCAAAATCACTTCAGTGTACATGAGTGAC 1392
 Db 456 HisLysPheLysHisArgArgGluAspPheCysSerAlaileGluCysTyTrpMetGluGlu 475
 QY 1393 TACAATGTCATCGGAGCGGAGGCG-----CGGAAGCACGTAATCGGTGATA 1440
 Db 476 TyrGlyValSerAlaGlnGluAlaTyAspValPheAsnLysHisValGlu----- 492
 QY 1441 GCGAGGTGTGGAAGAGATGATGATGCGAGAGGCTGCGAAGGATTCCTCCATTCGCAAA 1500
 Db 493 ---SerAlaTrpLysAspValAsnGlnGlu---PheGlnLysProThrGluMetProThr 510
 QY 1501 GATTTTATAGGATGTGAGTGTGATTTAGGAAGGATGCGCAGTGTGATGATCAATATGGA 1560
 Db 511 GluValLeuAsnArgSerLeuAsnLeuAlaArgValMetAspValLeuTyArgGluGly 530
 QY 1561 GATGGCAGCGCACAAACCTTATTATACATCAACAATGACCAAGAACCTTATTCGAG 1620
 Db 531 AspGlyTyTrpTyTrpValGlyLysAlaAlaLysGlyGlyIleThrSerLeuLeuileGlu 550
 QY 1621 CCCTTTTCCA 1629
 Db 551 ProIleAla 553

RESULT 4

DCS1_GOSAR STANDARD; PRT; 554 AA.
 AC Q39761.
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE (+)-delta-cadinene synthase isozyme XC1 (EC 4.2.3.13) (D-cadinene synthase).
 OS Gossypium arboreum (Tree cotton).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.

NCBI_TaxID=29729;
 RN [1] SEQUENCE FROM N.A.
 RC STRAIN=cv. Nanking;
 RX MEDLINE=96132653; PubMed=8554317;
 RA Chen X.-Y., Chen Y., Heinsteins P., Davison V.J.;
 RT "Cloning, expression, and characterization of (+)-delta-cadinene
 RL synthase: a catalyst for cotton phytoalexin biosynthesis.";
 RL Arch. Biochem. Biophys. 324:255-266(1995).
 CC -!- FUNCTION: Responsible for the cyclization of trans,trans-farnesyl
 CC diphosphate (FPP) to (+)-delta-cadinene.
 CC -!- CATALYTIC ACTIVITY: 2-trans,6-trans-farnesyl diphosphate = (+)-
 CC delta-cadinene + diphosphate.
 CC -!- PATHWAY: Phytoalexins gossypol and lacinilene C biosynthesis;
 CC first (committed) step.
 CC -!- SIMILARITY: Belongs to the terpene synthase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U23206; AA93064.1; -.
 DR PIR; S68365; S68365.
 DR HSP; Q40577; SEAU.
 DR InterPro; IPR008930; Terp_cyc_toroid.
 DR InterPro; IPR001906; Terp_synth-like.
 DR InterPro; IPR005630; Terpene_synth_C.
 DR InterPro; IPR008949; Terpenoid_synth.
 DR Pfam; PF01397; Terpene_synth_1.
 DR Pfam; PF03936; Terpene_synth_C; 1.
 DR Lysase; Multigene family.
 FT ACT_SITE 451 451 BY SIMILARITY.
 FT ACT_SITE 527 527 BY SIMILARITY.
 FT ACT_SITE 531 531 BY SIMILARITY.
 FT ACT_SITE 554 554 BY SIMILARITY.
 SQ SEQUENCE 554 AA; 64137 MW; 59D6922DED9DCAF CRC64;
 Alignment Scores:
 Pred. No.: 6,19e-49 Length: 554
 Score: 776.00 Matches: 178
 Percent Similarity: 52.07% Conservative: 111
 Best Local Similarity: 32.07% Mismatches: 232
 Query Match: 26.69% Indels: 34
 DB: 1 Gaps: 7
 US-09-938-956-6 (1-1632) x DCS1_GOSAR (1-554)
 QY 1 ATGAGACATCGGAACTACACCTTCCTCGTGG----- 36
 Db 21 MetArgProLysAlaAspPheGlnProSerIleTrpGlyAspLeuAenCysPro 40
 QY 37 GATGTCAACTTCATCCCAATCGCTTCTCAGTACTATAAGGAGGACAAACACGCTGATTAGG 96
 Db 41 AspLysAenIleAspAlaGluThrGluLysArgHisGlnGlnLeuLysGluValArg 60
 QY 97 GCTTCTGAGCTGCTCATTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 156
 Db 61 -----LysMetIleValAlaProMetAlaAenSerThrGln 72
 QY 157 CAACTTGAAGTTCATCGATGATCGACGATGATCGACGATGATCGACGATGATCGACGATGAT 216
 Db 73 LysLeuAlaPheIleAspSerValGlnArgLeuGlyValSerThrHisPheThrLysGlu 92
 QY 217 TTCAGAAATCTGTCTCTATATATCTGACCATCATCATATATACAGAACCTTTTCCA 276
 Db 93 IleGluAspGluLeuGluAenIleThrHisAenAen----- 105
 QY 277 AAAGAGAAAGGAGTCTCTACTCAGATCTCTGCAATTTAGGCTCTCAGAGAAATGCTGT 336
 Db 106 AspAlaGluAenAspLeuThrThrSerIleArgPheArgLeuLeuArgGluHisGly 125
 QY 337 TTTCAAGTCGCACAAAGAGGTATTGATGATGATTTCAAGAACGAGGAGGTGATTTCAAGAA 396
 Db 126 TyrAenValSerCysAspValPheAenLysPheLysAspGluGlnGlyAenPheLysSer 145
 QY 397 AGCTTTAGCAGCAGCAGCAGCAGGATTTGTCACATGATGATGATGATGATGATGATGATGAT 456
 Db 146 SerValThrSerAspValArgGlyLeuLeuGluLeuThrGlnAlaSerThrLeuArgVal 165
 QY 457 GAAGCGAAACCCAGCTCGAGTCAGCGAGGAGGAAATCCGACCAAAATTTTGGAGGAAAA 516
 Db 166 HisGlyGluAspIleLeuAspGluAlaIleSerPheThrThrHisHisLeuSerLeuAla 185
 QY 517 GTGAACGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 576
 Db 186 Val-----AlaSerLeuAspHisProLeuSerGluGluValSerHisAlaLeuLysGln 203
 QY 577 CCTCTTCATTGGAGGATTAAGCCCAATGATGATGATGATGATGATGATGATGATGATGATGAT 636
 Db 204 SerIleArgArgGlyLeuProArgValGluAlaArgHisThrLeuSerValThrGlnAsp 223
 QY 637 AGGCCCGACATGATCCAGTAGTGTGGAGCTTCCATCTGACCTAAATATTTGTTCAA 696
 Db 224 IleGluSerHisAenLysAlaLeuLeuGluPheAlaLysIleAspPheAenMetLeuGln 243
 QY 697 GCACATTTCAAGAGAGCTCAAGAAATCTTCAGTGTGTGGAGAAATATCTGGTGTGTT 756
 Db 244 PheLeuHisArgLysGluLeuSerGluIleCysArgTrpTrpLysAspLeuAspPheGln 263
 QY 757 GAGAGCTGCTCCTTCGCAAGGATAGTGTGTGGAGTGTGATGATGATGATGATGATGATGATGAT 816
 Db 264 ArgLysLeuProThrAlaArgAspArgValGluGlyThrPheTrpIleSerGlyVal 283
 QY 817 ATCAGGCCAGCTCAGCATGCAAGTCAAGATTAATGATGGCAAACTCAACCTCTGAT 876
 Db 284 TyrPheGluProGlnTyrSerLeuGlyArgLysMetLeuThrLysValIleAlaMetAla 303
 QY 877 ACGTGTATCATGATATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 936
 Db 304 SerIleValAspAspThrThrAspSerThrAlaThrThrGluGluLeuLeuProThrThr 323
 QY 937 GACCTCATTCGAGATGGGATATAACTCAATGACCAACTTCCGATTCAGATGCAACTG 996
 Db 324 AsnAlaIleGluArgTrpAspIleLysCysIleAspGluIleProGluThrMetLysPro 343
 QY 997 TGCTTTCTTCATCAACAACTTCGTCGATGATGATGATGATGATGATGATGATGATGATGAT 1056
 Db 344 SerTyrLysAlaLeuLeuAspValThrGluGluMetValGlnLeuValAlaGluHisGly 363
 QY 1057 GGCTCATCTTATACCTTACCTGCGGCAATCGTGGTGTGATGATGATGATGATGATGATGATGAT 1116
 Db 364 ArgGlnTyrArgValGluThrAlaLysAenAlaMetIleArgLeuAlaGlnSerThrLeu 383
 QY 1117 GTAGAGCAGCGTGTCTTACGCGCGGCACAAACCAAGTTTGGAGAGATATTGGAGAAC 1176
 Db 384 ValGluAlaLysTrpThrLeuGlnAenTyrLysProSerPheGluGluPheLysAlaAen 403
 QY 1177 TCATGCGAGTCGATAGTGGCCCTGTATGTTAAACGACATATTTCTCGAGTAACAGAT 1236
 Db 404 AlaLeuProThrCysGlyTyrAlaMetLeuAlaIleThrSerPheValGlyMetGlyAsp 423
 QY 1237 TCGTTTCAAGAGAGACCGTCGACAGTGTGTGACAAATACACAGATTTAGTTCCTGGTCA 1296
 Db 424 IleValThrProGluThrPheLysTrpAlaAlaSerAspProLysIleIleGlnAlaSer 443
 QY 1297 TCCTTCGTTCTGGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1356
 Db 444 ThrIleLeuCysArgPheMetAspValAlaGluHisLysPheLysHisArgArgGlu 463
 QY 1357 GATGTCGCGAAATCACTTCAGTGTACATGATGATGATGATGATGATGATGATGATGATGATGAT 1416
 Db 464 AspAspCysSerAlaIleGluCysTyrMetGluGluTyrGlyValThrAlaGlnGluAla 483

QY 1417 -----CGAAGCAGCGTGAATGCTCATAGCGGAGGTGTGGAAGAGATGAAT 1464
 Db 484 TyrAspValPheAsnLysHisValGlu-----SerAlaTrpLysAspLeuAsn 499
 QY 1465 GCGAGAGGCTGCGAAGGATTCTCCATTCGCGCAAGATTTTATAGGATGTCAGTTGAT 1524
 Db 500 GlnGluPheLeu---LysProThrGluMetProThrGluValLeuAsnArgSerLeuAsn 518
 QY 1525 TTAGGAGAGATGGCGAGTGTGATGATACCATATATGAGATGGCGACCGGCACACACCT 1584
 Db 519 LeuAlaArgValMetAspValLeuLeuTyrArgGluGlyAspGlyTyrThrTyrValGlyLys 538
 QY 1585 ATTATACATCAACAATGACCGAGAACCTTATTCGAGCCCTTTGCA 1629
 Db 539 AlaAlaLysGlyClyLeuThrSerLeuLeuLeuGluProIleAla 553
 RESULT 5
 DCSI_GOSHI
 AC P9365; STANDARD; PRT; 554 AA.
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE (+)-delta-cadinene synthase (EC 4.2.3.13) (D-cadinene synthase).
 GN CDNI.
 OS Gossypium hirsutum (Upland cotton).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosoids II; Malvales; Malvaceae; Malvoideae; Gossypium.
 OX NCBI_TaxID:3635;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Westburn M;
 RA Davis E.M., Chen Y.-S., Essenberg M., Pierce M.L.;
 RT "cDNA sequence of a (+)-delta-cadinene synthase gene induced in
 RT Gossypium hirsutum L. by bacterial infection.";
 RL [In] Plant Gene Register PGR98-040.
 RN [2]
 RP PARTIAL SEQUENCE, AND CHARACTERIZATION.
 RX MEDLINE=96351891; PubMed=8728715;
 RA Davis E.M., Tsuji J., Davis G.D., Pierce M.L., Essenberg M.;
 RT "Purification of (+)-delta-cadinene synthase, a sesquiterpene cyclase
 RT from bacteria-inoculated cotton foliar tissue.";
 RL Phytochemistry 41:1047-1055 (1996).
 CC -I- FUNCTION: Responsible for the cyclization of trans,trans-farnesyl
 CC diphosphate (FPP) to (+)-delta cadinene.
 CC -I- CATALYTIC ACTIVITY: 2-trans,6-trans-farnesyl diphosphate = (+)-
 CC delta-cadinene + diphosphate.
 CC -I- PATHWAY: Phytoalexins gossypol and lacinilene C biosynthesis;
 CC first (committed) step.
 CC -I- INDUCTION: By bacterial infection.
 CC -I- SIMILARITY: Belongs to the terpene synthase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U98318; AAC12784.1; --
 CC HSP; Q40577; SEAU
 DR InterPro; IPR008930; Terp_cyc_toroid.
 DR InterPro; IPR001906; Terp_synth-like.
 DR InterPro; IPR005630; Terpene synth C.
 DR InterPro; IPR008949; Terpenoid synth.
 DR Pfam; PF01397; Terpene synth; 1.
 DR Pfam; PF03936; Terpene_synth_C; 1.
 KW Lyase.
 FT ACT_SITE 451 BY SIMILARITY.
 FT ACT_SITE 527 BY SIMILARITY.
 FT ACT_SITE 531 BY SIMILARITY.

SQ SEQUENCE 554 AA; 64019 MW; 8BCC78AD8CA5B816 CRC64;
 Alignment Scores: 6.19e-49 Length: 554
 Pred. No.: 776.00 Matches: 181
 Score: 51.71% Conservative: 106
 Best Local Similarity: 32.61% Mismatches: 234
 Query Match: 26.69% Indels: 34
 DB: 1 Gaps: 7
 US-09-938-956-6 (1-1632) x DCSI_GOSHI (1-554)
 QY 1 ATGAGACGATCCGGAATACAACTCTCTCGTTGG----- 36
 Db 21 MetArgProLysAlaAspPheGlnProSerIleTrpGlyAspPhePheLeuAsnCysPro 40
 QY 37 GATGTCAACTTCATCCCAATCGCTTCTCAGTCAGTATATAGGAGGACACACGATAGG 96
 Db 41 AspLysAsnIleAspAlaGluThrGlnLysArgHisGlnGlnLeuLysGluValArg 60
 QY 97 GCTTCTGAGCTGGTCACTTTGGTGAAGATCGAACTCGAAGAAAGAAACGGATCAATTCGA 156
 Db 61 -----LysMetIleValAlaProMetAlaAsnSerThrLeu 72
 QY 157 CAATCTGAGTTGATCGATGCTTCGAGAGGATGGGCTGCTCCGATCATTTCCAAATAGAG 216
 Db 73 LysLeuAlaPheIleAspSerValGlnGlyLeuGlyValSerTyrHisPheThrLysGlu 92
 QY 217 TTCAAAGAAATCTTGTCTCTATATATCTGACCATCATATTACAGAAACCCCTTTTCCA 276
 Db 93 IleGluAspGluLeuGluAsnIleTyrHisAsnAsn----- 105
 QY 277 AAAGAAGAAAGGATCTCTACTCCACATCTCTTCGATTTAGGCTCTCAGAGAACATGGT 336
 Db 106 AspAlaGluAsnAspLeuTyrThrThrSerLeuArgPheArgLeuLeuArgGluHisGly 125
 QY 337 TTCAAGTCCGACNAGAGGTATTCGATAGTTTCAGAACGAGGAGGTGAGTTCAAAGAA 396
 Db 126 PheHisValSerCysAspValPheAsnLysPheLysAspGluGlnGlnLysPheSer 145
 QY 397 AGCTTTAGCAGCACACGAGGATTTGTCACACTGTATGAAGCTTCTTTCTGTGACG 456
 Db 146 SerValThrSerAspValArgGlyLeuLeuGluLeuTyrGlnAlaSerTyrLeuArgVal 165
 QY 457 GAAGCGAAACCGCTCGAGTCGAGGAGGATTCGCCACCAATTTTGGAGGAAAAA 516
 Db 166 HisGlyGluAspIleLeuAspGluAlaIleSerPheThrSerAsnHisLeuSerLeuAla 185
 QY 517 GTGAACGAGGCTGTTGATGGCGACCTTTTAAACAAGAAATCGCATATTCTTTGGACATC 576
 Db 186 Val-----AlaSerLeuAspHisProLeuSerGluLysValSerHisAlaLeuLysGln 203
 QY 577 CCTCTTCATTGGAGGATTAAGGCCAAATGCGACCTGTGTGGATCGAATGGTATAGGAAG 636
 Db 204 SerIleArgArgGlyLeuProArgValGluAlaArgHisTyrLeuSerValTyrGlnAsp 223
 QY 637 AGGCCCGACATGATCCAGTAGTGTGGAGCTTGCCATCTCGCATTAATATTTGTTCAA 696
 Db 224 IleGluSerHisAsnLysValLeuLeuGluPheAlaLysIleAspPheAsnMetValGln 243
 QY 697 GCACAATTTCAAGAAGAGCTCAAGAATCCTTCAGGTGGTGGAGAAATATCGGTTTGT 756
 Db 244 LeuLeuHisArgLysGluLeuSerGluIleSerArgTrpTrpLysAspLeuAspPheGln 263
 QY 757 GAGAGAGTCCCTCGCAAGGATAGACTGGTGGATGCTACTTTTGGATATCTGGGATC 816
 Db 264 ArgLysLeuProTyrAlaArgAspArgValValGluGlyTyrPheTrpIleSerGlyVal 283
 QY 817 ATCGAGCCAGCTCAGCATGCAAGTGCAGGCAAGGATATGATGGCGCAAGAGTCAACGCTCTGATT 876
 Db 284 TyrPheGluProGlnTyrSerLeuGlyArgLysMetLeuThrLysValIleAlaMetAla 303
 QY 877 ACGTGTATCATGATATTTATGATGTCTATGGCACCTTAGAAGACTCGAACATTCAC 936

```

Db 304 SerileValAspThrTyrAspSerTyrAlaThrTyrGluGluLeuLeuProTyrThr 323
QY 937 GACCTCATTCGAGATGGATATAAATCAATCGACCAACTTCCCGATTACATGCAACTG 996
Db 324 AsnAlaIleGluArgTrpAspIleLysCysIleAspGluLeuProGluTyrMetLysPro 343
QY 997 TGTCTTCTTGCACCTCAACACTTCGTCGATGATACATCGATGATGATGAGGAGAA 1056
Db 344 SerTyrLysAlaLeuLeuAspValTyrGluGluMetGluGluLeuValAlaGluHisGly 363
QY 1057 GCGGTCAAGCTTACCTACCTCGCGCAATCGTGGTGTGATTTGGCGGATAGTATATG 1116
Db 364 ArgGluTyrArgValGluTyrAlaLysAsnAlaMetIleArgLeuAlaGlnSerTyrLeu 383
QY 1117 GTAGAGCAGCGTGGTCTTACGGCGGGCAACACCAAGTTTGGAGAGTATTGGAGAAC 1176
Db 384 ValGluAlaArgTrpThrLeuGlnAsnTyrLysProSerPheGluGluPheLysAlaAsn 403
QY 1177 TCATGGCAGTCGATAGTGGCGCTGTATGTTAAACGCACATATCTTCGAGTAAACAGAT 1236
Db 404 AlaLeuProThrCysGlyTyrAlaMetLeuAlaIleThrSerPheValGlyMetGlyAsp 423
QY 1237 TCCTTCAAAAGAGAGACCGTCGACGATTTGTACAAATACACGATTTAGTCTGTTGGTCA 1296
Db 424 IleValThrProGluThrPheLysTrpAlaAlaAsnAspProLysIleIleGlnAlaSer 443
QY 1297 TCCTTCTGTTCTGGCTGTGATGATTTGGGAACCTCGTGGAGAGGTGAGCAGAGGG 1356
Db 444 ThrIleLeuCysArgPheMetAspValThrGluHisLysPheLysHisArgArgGlu 463
QY 1357 GATGTCGCGAAATCACTTCAGTCTGCTACATGAGTACTCAATGCAATCGATCGGAGCGGAGGG 1416
Db 464 AspAspCysSerAlaIleGluCysTyrMetGluGluTyrGlyValThrAlaGlnGluAla 483
QY 1417 -----CGGAGCAGCGTGAATGCTGATAGCGGAGGTGGGAGAGATGAT 1464
Db 484 TyrAspValPheAsnLysHisValGlu-----SerAlaTrpLysAspValAsn 499
QY 1465 GCGGAGGAGGTGTGCGAAGGATCTCCATTCGCGCAAGATTTTATAGGATGTCAGTGTAT 1524
Db 500 ---GlnGlyPheLeuLysProThrGluMetProThrGluValLeuAsnArgSerLeuAsn 518
QY 1525 TTAGGAGGATGGCGGATGATGATACCATATGAGAGATGGCAGCGCACACACCT 1584
Db 519 LeuAlaArgValMetAspValLeuTyrArgGluGlyAspGlyTyrThrTyrValGlyLys 538
QY 1585 ATTATACATCAACAAATGACAGAACTTATTCGAGCCCTTTGCA 1629
Db 539 AlaAlaLysGlyGlyIleThrSerLeuLeuIleGluProIleAla 553

RESULT 6
DCS3 GOSAR . STANDARD; PRT; 555 AA.
ID DCS3 GOSAR AC Q43714;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE (+)-delta-cadinene synthase isozyme A (EC 4.2.3.13) (D-cadinene synthase).
DE NCBI TaxID=29729;
OS Gossypium arboreum (Tree cotton).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
OX NCBI TaxID=29729;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nanking;
RX MEDLINE=97060798; PubMed=8904844;
RA Chen X.-Y., Wang M., Chen Y., Davison V.J., Heinstein P.;
RT "Cloning and heterologous expression of a second (+)-delta-cadinene synthase from Gossypium arboreum.";
RT

```

```

J. Nat. Prod. 59:944-951 (1996).
-|- FUNCTION: Responsible for the cyclization of trans,trans-farnesyl diphosphate (FPP) to (+)-delta-cadinene.
-|- CATALYTIC ACTIVITY: 2-trans,6-trans-farnesyl diphosphate = (+)-delta-cadinene + diphosphate.
-|- PATHWAY: Phytoalexins biosynthesis; laccinilene C biosynthesis; first (committed) step.
-|- SIMILARITY: Belongs to the terpene synthase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; X96429; CAA65289.1; -.
EMBL; U27535; AAB41259.1; -.
HSSP; Q40577; SEAT.
InterPro; IPR008930; Terp_cyc_toroid.
InterPro; IPR001906; Terp_synth-like.
InterPro; IPR005630; Terpene_synth_C.
InterPro; IPR008949; Terpenoid_synth.
Pfam; PF01397; Terpene_synth_1.
Pfam; PF03936; Terpene_synth_C; 1.
Lyase; Multigene family.
ACT_SITE 452 452 BY SIMILARITY.
ACT_SITE 528 528 BY SIMILARITY.
ACT_SITE 532 532 BY SIMILARITY.
SEQUENCE 555 AA; 64103 MW; 7060C4F9D99412F9 CRC64;

Alignment Scores:
Pred. NO.: 2,62e-46 Length: 555
Score: 740.00 Matches: 174
Percent Similarity: 51.91% Conservative: 111
Best Local Similarity: 31.69% Mismatches: 238
Query Match: 25.45% Indels: 26
DB: 1 Gaps: 7

US-09-938-956-6 (1-1632) x DCS3_GOSAR (1-555)
QY 4 AGACGATCGCGAACTACAACCTTCCTCGTTGGGATGTCACACTTCATCCAAATCGCTTCTC 63
Db 21 ArgProLysAlaAspPheHisProGlyTyrGlyAspMetPheIleLeuCysProAsp 40
QY 64 AGTGAC-----TATAAGGAGGACAAACCGTATTAGGCT 99
Db 41 ThrAspIleAspAlaAlaThrGluLeuGlnTyrGluLeuLysAlaGlnValArg--- 59
QY 100 TCTGAGCTGTCACCTTTGGTGAAGATGAACTGGAGAAAGAAACCGATCAAATTCGACAA 159
Db 60 -----LysMetIleMetGluProValAspAspSerAsnGlnLys 72
QY 160 CTTGAGTTGATCGATGATGTCAGAGATGGGCTGTCGATCATTTCCAAATAGAGTTC 219
Db 73 LeuProPheIleAspAlaValGlnArgLeuGlyValSerTyrHisPheGluLysGluIle 92
QY 220 AAAGAAATCTTGCTCTATATATCTCGACCATCATATTACAGAACCCCTTTTCCAAA 279
Db 93 GluAspGluLeuGluAsnIleTyrArgAspThrAsnAsn-----AsnAsp 107
QY 280 GAAGAAGAGGATCTTACTCCACATCTCTTGCAATTTAGGCTCTCAGAGAACATCGTTT 339
Db 108 AlaAspThrAspLeuTyrThrThrAlaLeuArgPheArgLeuLeuArgGluHisGlyPhe 127
QY 340 CAATCGCACAGAGGATTTTCGATAGTTCAGAACCGAGGAGGTGATTCAAAGAAAGC 399
Db 128 AspIleSerCysAspAlaPheAsnLysPheLysAspGluAlaGlyAsnPheLysAlaSer 147
QY 400 CTTAGCAGCAGCACACGAGGATTTGTCACATGTATGAGCTTCTTCTGTTGACCGAA 459
Db 148 LeuThrSerAspValGlnGlyLeuLeuGluLeuTyrGluAlaSerTyrMetArgValHis 167

```

QY	460	GGCGAAACACGCTCGAGTCAGCGAGGAATTTCCCHACCAAAATTTTGGAGGAAAAAGTG	519
Db	168	GlyGluAspLeuAspGluAlaIleSerPheThrThrAlaGlnLeuThrLeuAlaLeu	187
QY	520	AACGAGGGTGGTGTGATGGCGACCTTTTAAACAAGAAATCGCATATCTTTTGGACATCCCT	579
Db	188	Prothr-----LeuHisHisProLeuSerGluGlnValGlyHisAlaLeuLysGlnSer	205
QY	580	CTTCATTGGAGGATTAAGACCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGG	639
Db	206	IleArgArgGlyLeuProArgValGluAlaArgAsnPheIleSerIleTyrGlnAspLeu	225
QY	640	CCCGACATGAATCCAGTAGTGTGGAGCTTGCATACTCGACATTAATATGTGTCAAGCA	699
Db	226	GluSerHisAsnLysSerLeuLeuGlnPheAlaLysIleAspPheAsnLeuLeuGlnLeu	245
QY	700	CAATTTCAAGAGACTCAAGAAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTGAG	759
Db	246	LeuHisArgLysGluLeuSerGluIleCysArgTrpTrpLysAspLeuAspPheThrArg	265
QY	760	AAGTCGCCCTTCGCAAGGATAGACTGCTGGAATGCTACTTTTGGATATCTGGGATCATC	819
Db	266	LysLeuProPheAlaArgAspArgValValGluGlyTyrPheTrpIleMetGlyValTyr	285
QY	820	GAGCCAGCTCAGCATGCAAGTCGCAAGATAATGATGGCGCAAGTCAACGCTCTGATTACG	879
Db	286	PheGluProGlnTyrSerLeuGlyArgLysMetLeuThrLysValIleAlaMetAlaSer	305
QY	880	GTGATCATGATATTTATGATGTCATGCGCACCTTAGAAGAACTCGAACAATTCACCTGAC	939
Db	306	IleValAspAspThrTyrAspSerTyrAlaThrTyrAspGluLeuIleProTyrThrAsn	325
QY	940	CTCATTCGAAGATGGGTATAAACTCAATGCCACCACTCCCGATTACATCGCAACTGTGC	999
Db	326	AlaIleGluArgTrpAspIleLysCysMetAsnGlnLeuProAsnTyrMetLysIleSer	345
QY	1000	TTTCTTCGACTCAACACTTCGTCGATGATACATCGTACGATGTTAGAGGAGAAGGC	1059
Db	346	TyrLysAlaLeuLeuAsnValTyrGluGluMetGluGlnLeuLeuAlaAsnGlnGlyArg	365
QY	1060	GTCAACGTTATACCTACCTCGCGCAATCGTGGGTTGATTTGGCGGATAAGTATATGGTA	1119
Db	366	GlnTyrArgValGluTyrAlaLysLysAlaMetIleArgLeuValGlnAlaTyrLeuLeu	385
QY	1120	GAGCAGGTGGTCTACGGGGGCGCAACAACCAAGTTTGGAGAGTATTTGGAGAACTCA	1179
Db	386	GluAlaLysTrpThrHisGlnAsnTyrLysProThrPheGluGluPheArgAspAsnAla	405
QY	1180	TGGCAGTCGATAGTGGGCGCTGTATGTTAAACG---CACATATCTTCCAGTAAACAGAT	1236
Db	406	LeuProthr---SerGlyTyrAlaMetLeuAlaIleThrAlaPheValGlyMetGlyGlu	424
QY	1237	TCGTTTCAAGAGGACCGTCGACAGTTTGTACAAATACACAGATTTAGTTCGTTGGTCA	1296
Db	425	ValIleThrProGluThrPheLysTrpAlaAlaSerAspProLysIleIleLysAlaSer	444
QY	1297	TCCTTCGTTCTCGCGCTGTGTGATGATTTGGGAACCTCCGTCGAAGAGGTGAGCAGAGG	1356
Db	445	ThrIleLysArgPheMetAspAspIleAlaGluHisLysPheAsnHisArgArgGlu	464
QY	1357	GATGTCCGAAATCACTTCAGTGCTCATGAGTGACTACAAATGCATCGAGGCGGAGCGG	1416
Db	465	AspAspCysSerAlaIleGluCysTyrMetLysGlnTyrGlyValThrAlaGlnGluAla	484
QY	1417	CGGAAGCAGCTGAAATGCCTGATAGCGGAGTGTGGAAAGAGATGAATCGGAGAGGGTG	1476
Db	485	TyrAsnGluPheAsnLysHisIleGluSerSerTrpLysAspValAsn---GluGluPhe	503
QY	1477	TCGAAGGATTCTCCATTTCCGCAAGATATTTATAGGATGTGCAGTTGATTTAGGAAGATG	1536
Db	504	LeuLysProThrGluMetProThrProValLeuCysArgSerLeuAsnLeuAlaArgVal	523

```

Qy 1537 GCGCAGTTGATGTACCATATGAGATGGCAGGCAGGCACACACACCTTATTACATCAA 1539
Db 524 MetAspValLeuYrArgGluGlyAspGlyTyrThrHisValGlyLysAlaLysGly 543
Qy 1597 CAAATGACCAAGAACCTTATTTCGAGCCC 1623
Db 544 GlyIleThrSerLeuLeuIleAspPro 552

RESULT 7
TSDI_ABIGR STANDARD; PRT; 628 AA.
ID TSD1_ABIGR AC O24475;
DT 28-FEB-2003 (Rel. 41, Last Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Pinene synthase, chloroplast precursor (EC 4.2.3.14) (Beta-geraniolene
DE synthase) ((-)-(1S,5S)-pinene synthase).
GN AG3.18.
OS Abies grandis (Grand fir).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.
CN NCBI_TaxID=46611;
RX [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=97413772; PubMed=9268308;
RA Boehlmann J., Steele C.L., Croteau R.;
RT "Monoterpene synthases from grand fir (Abies grandis). cDNA isolation,
RT characterization, and functional expression of myrcene synthase, (-)-
RT (4S)-limonene synthase, and (-)-(1S,5S)-pinene synthase.";
RL J. Biol. Chem. 272:21784-21792(1997).
CC -1- FUNCTION: Involved in defensive oleoresin formation in conifers in
CC response to insect attack or other injury. Involved in monoterpene
CC (C10) olefins biosynthesis. A mixture of alpha- and beta-pinene is
CC produced by this enzyme.
CC -1- CATALYTIC ACTIVITY: Geranyl diphosphate = pinene + diphosphate.
CC -1- COFACTOR: Manganese and potassium.
CC -1- PATHWAY: Oleoresinosis.
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- INDUCTION: By wounding.
CC -1- SIMILARITY: Belongs to the terpene synthase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U87909; AAB71085.1; -.
CC HSPG; Q40577; SEAT.
CC InterPro; IPR008930; Terp_cyc_toroid.
CC InterPro; IPR001906; Terp_synth-like.
CC InterPro; IPR005630; Terpene_synth_C.
CC InterPro; IPR008949; Terpenoid_synth.
CC Pfam; PF01397; Terpene_synth_1.
CC Pfam; PF03936; Terpene_synth_C; 1.
CC Lyase; Manganese; Transit peptide; Chloroplast.
CC TRANSIT 1 ? CHLOROPLAST (POTENTIAL).
CC CHAIN 2 628 PINENE SYNTHASE.
CC ACT_SITE 524 524 BY SIMILARITY.
CC ACT_SITE 601 601 BY SIMILARITY.
CC ACT_SITE 605 605 BY SIMILARITY.
CC
CC SEQUENCE 628 AA; 71505 MW; 23DBB78BF9C8072C CRC64;

Alignment Scores:
Pred. No.: 2,12e-44 Length: 628
Score: 714.00 Matches: 191
Percent Similarity: 52.19% Conservative: 107
Best Local Similarity: 33.45% Mismatches: 231
Query Match: 42 Indels: 42
DB: 1 Gaps: 14

```


CC -!- FUNCTION: Catalyzes the cyclization of geranylgeranyl diphosphate
CC to casbene, a diterpene phytoalexin with antibacterial and
CC antifungal activity.
CC -!- CATALYTIC ACTIVITY: Geranylgeranyl diphosphate = casbene +
CC diphosphate.
CC -!- SUBCELLULAR LOCATION: Chloroplast.
CC -!- INDUCTION: By oligogalacturonide fragments released by fungal
CC infection. Detected after 5 h of incubation with the pectic
CC fragments and reaches a maximum after 10-12 h.
CC -!- MISCELLANEOUS: The Km of this enzyme is 1.9 micromol.
CC -!- SIMILARITY: Belongs to the terpene synthase family.
CC -!- CAUTION: It is uncertain whether Met-1 or Met-8 is the initiator.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L32134; -; NOT ANNOTATED CDS.
CC InterPro; IPR008930; Terp_cyc_coreid.
CC InterPro; IPR001906; Terp_synth-like.
CC InterPro; IPR005630; Terpene synth C.
CC InterPro; IPR008949; Terpenoid synth.
CC Pfam; PF01397; Terpene synth; I.
CC Pfam; PF03936; Terpene synth; C.1.
CC Plant defense; Lyase; Chloroplast; Transit peptide.
CC TRANSIT 1 56 CHLOROPLAST (POTENTIAL).
CC CHAIN 57 601 CASBENE SYNTHASE.
CC ACT_SITE 500 500 BY SIMILARITY.
CC ACT_SITE 575 575 BY SIMILARITY.
CC ACT_SITE 579 579 BY SIMILARITY.
CC SEQUENCE 601 AA; 68965 MW; F7B362D286747957 CRC64;
Alignment Scores:
Pred. No.: 3,656-43 Length: 601
Score: 697.00 Matches: 167
Percent Similarity: 51.21% Conservative: 109
Best Local Similarity: 30.98% Mismatches: 245
Query Match: 23.97% Indels: 18
Dbs: 1 Gaps: 7
US-09-938-956-6 (1-1632) x CASS_RICCO (1-601)
QY 19 TACAACCTTCGTTGGGATGTCACACTTCATCCCAATCGCTCTCAGTGACTATAAGGAG 78
Db 69 PheProProThrValTrpGlyAsnArgPheAlaSerLeuThrPheAsnProSerGluPhe 88
QY 79 GACAAACACGATGATTAGGCTTCTGAGCTGTGCTACCTTTGGTGAG---ATGGAACTGGAG 135
Db 89 GluSerTyrAspGluArgValIleValLeuLysLysValLysAspIleLeuSer 108
QY 136 AAGAAACGGATCAAAATTCGACAACTTCGATTGATCGATGACTTCGACGAGATGGGCTG 195
Db 109 SerThrSerAspSerValGluThrValIleLeuLeuAspLeuLeuCysArgLeuGlyVal 128
QY 196 TCCGATCATTTCCAAATAGTTCAAAGAAATCTTGTCTCTATATATCTCGACCATCAC 255
Db 129 SerTyrHisPheGluAsnAspIleGluLeuLeuSerLysIlePhe----- 144
QY 256 TATTACAGAACCTTTTCCA-----AAGAAAGAGGATCTCTACTCCACA 303
Db 145 -----AsnSerGlnProAspLeuValAspGluLysGluCysAspLeuThrAla 161
QY 304 TCTCTTGCAATTTAGGCTCTCAGAGAACATGTTTTCAAGTCGCAAGAGGATTTCCGAT 363
Db 162 AlaIleValPheArgValPheArgGlnHisGlyPheLysMetSerSerAspValPheSer 181
QY 364 ACTTTCAAGACGAGGCTGAGTTCAAAGAACCTTAGCGACACACACAGGATTG 423
Db 182 LysPheLysAspSerAspGlyLysPheLysGluSerLysGluArgGlyAspAlaLysGlyMet 201

QY 424 TTCCAACCTGTATGACGCTTCCTTTCTGTGTGACGGAAGCGGAACACCGCTCGAGTCAGCG 483
Db 202 LeuSerLeuPheGluAlaSerHisLeuSerValHisGlyGluAspIleLeuGluGluAla 221
QY 484 AGGGAATTCGCCACCAAAATTTTGGAGGAAAGATGACAGAGGGTGTGTGTATGCGGAC 543
Db 222 PheAlaPheThrLysAspTyrLeuGlnSerSerAlaValGlu-----LeuPheProAsn 239
QY 544 CTTTTAACAGAAATCCATATCTTCTGGACATCCCTCTCTCATTTGGAGGATTTAAAGGCCA 603
Db 240 LeuLysArgHisIleThrAsnAlaLeuGluGlnProPheHisSerGlyValProArgLeu 259
QY 604 AATGCACCTGTGTGGATCGAATGTATAGGAAGAGCCCGAC---ATGAATCCAGTAGTG 660
Db 260 GluAlaArgLysPheIleAspLeuThrGluAlaAspIleGluCysArgAsnGluThrLeu 279
QY 661 TTGGAGCTTGGCCATCTCGACTTAAATATTTTCAAGCACAATTTCAAGAGAGACTCAA 720
Db 280 LeuGluPheAlaLysLeuAspTyrAsnArgValGlnLeuLeuHisGlnGluLeuCys 299
QY 721 GAATCCTTCAGGTGGTGGGAATACTGGTGTGTGTGAGAGCTGCCCTTCGCAAGGGAT 780
Db 300 GlnPheSerLysTrpTrpLysAspLeuAsnLeuAlaSerAspIleProTyrAlaArgAsp 319
QY 781 AGACTGTGTGAATGCTACTTTTGGATATCTGGATCATCGAGCCAGCTCAGCATGCAAGT 840
Db 320 ArgMetAlaGluIlePhePheTrpAlaValAlaMetTyrPheGluProAspTyrAlaHis 339
QY 841 GCAAGATATGATGGGCAAGTCAAGCTCTGATTACGGTGTGATGATGATGATGATGAT 900
Db 340 ThrArgMetIleIleAlaLysValValLeuLeuLeuSerLeuLeuAspThrIleAsp 359
QY 901 GTCTATGGCACCCTTAGAAGAACTCGAACTTCACTGACCTCATTCGAGATGGGATATA 960
Db 360 AlaTyrAlaThrMetGluGluThrHisIleLeuAlaGluAlaValAlaArgTrpAspMet 379
QY 961 AACTCAATCGACCAACTTCCGATACATGCAACTGCTGCTTCTTCCATCACTCAACTTC 1020
Db 380 SerCysLeuGluLysLeuProAspTyrMetLysValIleTyrLysLeuLeuAsnThr 399
QY 1021 GTCGATGATACATCGTACGATGTTATGAAGGAGAAAGGCGTCAACGTTATCCCTACCTG 1080
Db 400 PheSerGluPheGluLysGluLeuThrAlaGluLysSerTyrSerValLysTyrGly 419
QY 1081 CGCAATCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1140
Db 420 ArgGluAlaPheGlnGluLeuValArgGlyTyrTyrLeuGluAlaValTrpArgAspGlu 439
QY 1141 GGCACAAACCAAGTTTGGAGAGTATTGGAGAACTCATCGCAGTCGATAAGTGGGCC 1200
Db 440 GlyLysIleProSerPheAspAspTyrLeuTyrAsnGlySerMetThrThrGlyLeuPro 459
QY 1201 TGTATGTTAACGCACATATTTCTCCGAGTAACAGATTCGTTCCAAAGAGACCGTCGAC 1260
Db 460 LeuValSerThrAlaSerPheMetGlyValGlnGluIleThrGlyLeuAsnGluPheGln 479
QY 1261 AGTTTGTACAAATACCAAGTATTAGTTCTGTGTGTCATCTCTGCTCGGCTTGTGAT 1320
Db 480 TrpLeuGluThrAsnProLysLeuSerTyrAlaSerGlyAlaPheIleArgLeuValAsn 499
QY 1321 GATTTCGCAACCTCGGTGGAGAGGTCGAGCAGAGGGGATGTCCGAAATCATCTTCAGTGC 1380
Db 500 AspLeuThrSerHisValThrGluGlnArgGlyHisValAlaSerCysIleAspCys 519
QY 1381 TACATGATGACTACATGCTCGAGGCGGAGCGCGGAGCAGCAGTGAATCGCTGATA 1440
Db 520 TyrMetAsnGlnHisGlyValSerLysAspGluAlaValLysIleLeuGlnLysMetAla 539
QY 1441 GCGAGGTGTGGAAGAGATGATCGGAGAGGCTGTGCAAGGATTTCTCCATTCGCGCAA 1500
Db 540 ThrAspCysTrpLysGluIleAsn---GluGluCysMetArgGlnSerGlnValSerVal 558


```

Db      418 LeuProGluTyrMetLysGlyValTyrMetAspLeuTyrGlnCysIleAsnGluMetVal 437
Qy      1036 TACGATGTTATCAAGAGAGAGCGCTCAACGTTATACCTTACCTCCGCGCAATCGTGGTT 1095
Db      438 GluGluAlaGluLysThrGlnGlyArgAspMetLeuAsnTyrIleGlnAsnAlaTyrGlu 457
Qy      1096 GATTTCGGCGGATATATGATGATGACGACGCGGTCTACGCGCGGCGCAACACCAAGT 1155
Db      458 AlaLeuPheAspThrPheMetGlnGluAlaLysTyrIleSerSerSerTyrLeuProThr 477
Qy      1156 TTGGAAGAGTATTTCGAGAACTCATGCGCAGTCGATTAAGTGGCGCCCTGTATTAACGCAC 1215
Db      478 PheGluGluTyrLeuLysAsnAlaLysValSer---SerGlySerArgIleAlaThrLeu 496
Qy      1216 ATATTCTCCGAGTCAACAGAT-----TCGTTCAACAAGGACCGTCGACAGT 1263
Db      497 GlnProIleLeuThrLeuAspValProLeuProAspTyrIleLeuGlnGluIleAspTyr 516
Qy      1264 TTGTACAATACACAGATTAGTTCGTTGGTTCATCTCGTTCGTCGCGCTTCGCTGATGAT 1323
Db      517 ProSerArgPheAsnGluLeu-----AlaSerSerIleLeuArgLeuArgGlyAsp 533
Qy      1324 TTGGGAACCTCGGTGGAAGAGTGAGCAGAGGGGATGTCCGGAATCACTTCAGTGTCTAC 1383
Db      534 ThrArgCysTyrLysAlaAspArgAlaArgGlyGluGluAlaSerAlaIleSerCysTyr 553
Qy      1384 ATGAGTGTACTAATGCATCG---GAGCGGAGGCGCGAAGCAGTCGAAATGGCTGATA 1440
Db      554 MetLysAspHisProGlySerIleGluGluAspAlaLeuAsnHisIleAsnAlaMetIle 573
Qy      1441 GCGGAGGTGGAAGAGATGAATGCGGAGAGGGTGTGGAAGGATTCCTCAATTCGGCAAA 1500
Db      574 SerAspAlaIleArgGluLeuAsnTyrGluLeuLeuArgProAspSerLysSerProIle 593
Qy      1501 GATTTTATAGGATGTCAGTTGATTTAGGAAGGATGGCGCAGTTCATGATCAATATGGA 1560
Db      594 SerSerLysLysHisAlaPheAspIleThrArgAlaPheHisValTyrLysTyrArg 613
Qy      1561 GATGGCAGCGGCACACAAACCCCTATTATATACAACTATACAACTACCAAGACCTTATTCGAG 1620
Db      614 AspGlyTyrThrValSerAsnAsnGluThrLysAsnLeuValMetLysThrValLeuGlu 633
Qy      1621 CCTTTGCA 1629
Db      634 ProLeuAla 636

RESULT 10
TS2D_ABIGR STANDARD; PRT; 627 AA.
ID TS2D_ABIGR
AC 024474;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Myrcene synthase, chloroplast precursor [EC 4.2.3.15].
GN AG2.2.
OS Abies grandis (Grand fir).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.
OC NCBI_TaxID=46611;
RN [1]
RN SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=97413772; PubMed=9268308;
RA Bohlmann J., Steele C.L., Croteau R.;
RT "Monoterpene synthases from grand fir (Abies grandis). cDNA isolation,
RT characterization, and functional expression of myrcene synthase, (-)-
RT (4S)-limonene synthase, and (-)-(1S,5S)-pinene synthase."
RL J. Biol. Chem. 272:21784-21792(1997).
CC -1- FUNCTION: Involved in defensive oleoresin formation in conifers in
CC response to insect attack or other injury. Involved in monoterpene
CC (C10) olefins biosynthesis.
CC -1- CATALYTIC ACTIVITY: Geranyl diphosphate = myrcene + diphosphate.
CC -1- COFACTOR: Manganese and potassium.
CC -1- PATHWAY: Oleoresinosis.
```

```

QY 469 ACCTCGAGTCAGGAGGATTCGCCACCAATTTTCGAGGAAAAAGTGAACGGGT 528
Db 235 ValMetGluGluAlaGluLeuPheSerThrArgTyrLeuLysGluAlaLeuGluLysIle 254
QY 529 GGTGTGATGGCGACCTTTTAAACAAGATCGCATATTCCTTGGACATCCCTCTTCATTGG 588
Db 255 ProValSerAla---LeuSerGlnGluIleLysPheValMetGluTyrGlyTrpHisThr 273
QY 589 AGGATTAAAGGCAATGCACTGTGTGATCGAA----- 624
Db 274 AsnLeuProArgLeuGluAlaArgAsnTyrIleAspThrLeuGluLysAspThrSerAla 293
QY 625 TGTGTATAGGAGGCGCCGACATGATCCAGTGTGTGAGCTTCGCCATATCCGACTTA 684
Db 294 TrpLeuAsnLys-----AsnAlaGlyLysLeuLeuGluLeuAlaLysLeuGluPhe 311
QY 685 AATATTGTTCAAGCACATTTTCAAGAAGAGCTCAAGATCCTTCAGGTGGTGGAGAAAT 744
Db 312 AsnIlePheAsnSerLeuGlnGlnLysGluLeuGlnTyrLeuLeuArgTrpTrpLysGlu 331
QY 745 ACTGGGTTCTGTGAGAGCTGCCCTTCGCAAGGATAGACTGTGTGGAATGCTACTTTGG 804
Db 332 SerAsp---LeuProLysLeuThrPheAlaArgHisArgHisValGluPheTyrThrLeu 350
QY 805 AATACTGGGATCATCGGCCACCTCGACATGCAAGTGCAGATGATGATGATGATGATGATG 864
Db 351 AlaSerCysIleAlaIleAspProLysHisSerAlaPheArgLeuGlyPheAlaLysMet 370
QY 865 AACGCTGATTCAGGTGATCGATGATATTTATGATGCTATGCGACCTTAGAAGAACATC 924
Db 371 CysHisLeuValThrValLeuAspPheIleTyrAspThrPheLysThrIleAspGluLeu 390
QY 925 GAACAATTCATGCTCATTCGAGATGGATATAAATCAATCAATGCAACCACTCCGAT 984
Db 391 GluLeuPheThrSerAlaIleLysArgTrpAsnSerSerGluIleGluHisLeuProGlu 410
QY 985 TAGATGCACTGCTCTTCTTGCACTCAACCACTTCGTCATGATCATCATCGATCGATGTT 1044
Db 411 TyrMetLysCysValTyrMetValPheGluThrValAsnGluLeuThrArgGluAla 430
QY 1045 ATGAAGAGAAAGCGTCAACGTTATACCTCTCGCGCAATCGTGGGTGATTTGGCG 1104
Db 431 GluLysThrGlnGlyArgAsnThrLeuAsnTyrValArgLysAlaTrpGluAlaTyrPhe 450
QY 1105 GATAAGTATATGTAGAGGACGGTGTCTTACGGCGGCGACAAACCAAGTTGGAAGAG 1164
Db 451 AspSerTyrMetGluGluAlaLysTrpIleSerAsnGlyTyrLeuProMetPheGluGlu 470
QY 1165 TATTGGAGAACTCATGCGAGTCGATAGTGGGCGCTGTATGTTAAGCAGCATATCTTC 1224
Db 471 TyrHisGluAsnGlyLysValSerSerAla-----Tyr 481
QY 1225 CGAGTAACA-----GATTGTTCCACAAG 1248
Db 482 ArgValAlaThrLeuGlnProIleLeuThrLeuAsnAlaTrpLeuProAspTyrIleLeu 501
QY 1249 GAGACCGTCGACAGTTGTACAAATACCAATTTAGTTCGTTGGTTCATCTCGTCTCG 1308
Db 502 LysGlyIleAspPheProSerArgPheAsnAspLeuAla-----SerSerPhe---Leu 518
QY 1309 CGGCTTGTGATGATTTGGGAACCTCGGTGGAAGAGGTGACGAGAGGGATGTCGCGAAA 1368
Db 519 ArgLeuArgGlyAspThrArgCysTyrLysAlaAspArgAspArgGlyGluGluAlaSer 538
QY 1369 TCACCTTGTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1425
Db 539 CysIleSerCysTyrMetLysAspAsnProGlySerThrGluAlaAspAlaLeuAsnHis 558
QY 1426 GTGAAATGCTGATAGCGAGGTGTGGAAGAGATGAATGCGAGAGGGTGTGCAAGGAT 1485
Db 559 IleAsnAlaMetValAsnAspIleIleLysGluLeuAsnTrpGluLeuLeuArgSerAsn 578

```

```

QY 1486 TCTCCATTCCGCAAGATTTTATAGGATGTGCAATTTATAGGATGGCGAGGATGGCGAGTTG 1545
Db 579 AspAsnIleProMetLeuAlaLysLysHisAlaPheAspIleThrArgAlaLeuHisHis 598
QY 1546 ATGTACCATATGAGATGGCGGCGGACACACACACACACACACCTATTATACATCAACAAATGACC 1605
Db 599 LeuTyrIleTyrArgAspGlyPheSerValAlaAsnLysGluThrLysLysLeuValMet 618
QY 1606 AGAACCTTATTCGAG 1620
Db 619 GluThrLeuLeuGlu 623

```

RESULT 11

TASY TAXER STANDARD; PRT; 862 AA.

ID Q41594; Q94FV8; AC 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Taxadiene synthase (EC 4.2.3.17) (Taxa-4(5),11(12)-diene synthase).

GN TDCL.

OS *Taxus brevifolia* (Pacific yew).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Coniferopsida; Coniferales; Taxaceae; Taxus.

OX NCBI_TaxID=46220;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96199163; PubMed=8621577;

RT Wildung M.R., Croteau R.B.;

RT "A cDNA clone for taxadiene synthase, the diterpene cyclase that catalyzes the committed step of taxol biosynthesis.";

RL J. Biol. Chem. 271:9201-9204(1996).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=21297238; PubMed=11404343;

RT Trapp S.C., Croteau R.B.;

RT "Genomic organization of plant terpene synthases and molecular evolutionary implications";

RL Genetics 158:811-832(2001).

CC -!- FUNCTION: Catalyzes the cyclization of the ubiquitous isoprenoid intermediate geranylgeranyl diphosphate to taxa-4,11-diene, the parent olefin with a taxane skeleton.

CC -!- CATALYTIC ACTIVITY: Geranylgeranyl diphosphate = taxa-4,11-diene + diphosphate.

CC -!- PATHWAY: Taxol biosynthesis; first step.

CC -!- SIMILARITY: Belongs to the terpene synthase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; U48796; AAC49310.1; ..

CC EMBL; AF326519; AAK83566.1; ..

DR HSSP; Q40577; SEAU.

DR InterPro; IPR008930; Terp_cyc_toroid.

DR InterPro; IPR001906; Terp_synth-like.

DR InterPro; IPR005630; Terpene synth C.

DR InterPro; IPR008949; Terpenoid synth.

DR Pfam; PF01397; Terpene_synth; 1.

DR Pfam; PF03936; Terpene_synth_C; 1.

DR Taxol biosynthesis; Lyase.

KW ACT SITE 758 758 BY SIMILARITY.

FT ACT SITE 835 835 BY SIMILARITY.

FT ACT SITE 839 839 BY SIMILARITY.

FT CONFLICT 148 148 L -> V (IN REF. 2).

FT CONFLICT 767 767 A -> V (IN REF. 2).

SQ SEQUENCE 862 AA; 98303 MW; 9141BS9780CD79A1 CRC64;

Alignment Scores:

Pred. No.:	7,956-39	Length:	862
Score:	638.00	Matches:	161
Percent Similarity:	50.86%	Conservative:	106
Best Local Similarity:	30.67%	Mismatches:	210
Query Match:	21.94%	Indels:	48
DB:	1	Gaps:	17

US-09-938-956-6 (1-1632) x TASY_TAXCH (1-862)	
QY 145 GATCAAAATTCGACAACTTCAGTTGATCGATGATCTTCGACAGAGATGGGCTGTCGATCAT 204	
DB 351 AspLeuLeuGluArgLeuSerLeuValAspAsnIleGluHisLeuGlyIleGlyArgHis 370	
QY 205 TTCGAAATAGTTCGAAAGAAATCTGCTCTATATATCTCGACCATCATATTTACAAG 264	
DB 371 PheLysGlnGluLeuLysGlyAlaLeuAspTyrValTyr-----ArgHisTrp----- 386	
QY 265 AACCCCTTTTCCAAAAGAAAGG-----GAT 291	
DB 387 -----SerGluArgGlyIleGlyTyrGlyArgSerLeuValProAsp 401	
QY 292 CTCCTACTCCACATCTCTTCATTTAGCTCTCTCAGACACATGCTTTTCAGTGCACAA 351	
DB 402 LeuAsnThrThrAlaLeuGlyLeuArgThrLeuArgMetHisGlyTyrAsnValSerSer 421	
QY 352 GAGGTATTCGATGTTTCAAGAACAGAGAGGCTGAGTTCAAGAAAGCTTAGCCGAC--- 408	
DB 422 AspValLeuAsnAsnPheLysAspGluAsnGlyArgPhePheSerSerAlaGlyGlnThr 441	
QY 409 -----GACACACAGAGATTTGCACTGTATGAGCTTCTCTCTGTTGAGGAGGC 462	
DB 442 HisValGluLeuArgSerValValAsnLeuPheArgAlaSerAspLeuAlaPheProAsp 461	
QY 463 GAAACACGCTCGAGTCAGGAGGAATTCGCCACCAATTTTTCGAGGAAAGTG--- 519	
DB 462 GluArgAlaMetAspAspAlaArgLysPheAlaGluProTyrLeuArgGluAlaLeuAla 481	
QY 520 AACGAGGCTGTGTTGCGACGACCTTTTAAACAGAAATCGCATATCTTTGACATCCCT 579	
DB 482 ThrLysIleSerThrAsnThrLysLeuPheLysGluIleGluTyrValValGluTyrPro 501	
QY 580 CTCCTATGAGGNTTAAAGGCCAATGCACTGTGTGATCGAA----- 624	
DB 502 TrpHisMetSerIleProArgLeuGluAlaArgSerTyrIleAspSerTyrAspAspAsn 521	
QY 625 -----TGCTATAGGAAG-----AGCCCGACATG---AATCCAGTAGTGTG 663	
DB 522 TyrValTrpGlnArgLysThrLeuTyrArgMetProSerLeuSerAsnSerLysCysLeu 541	
QY 664 GAGCTTGCCATCTGCACTTAATATGTTTCAAGCACAATTTCAAGACAGAGCTCAAGAA 723	
DB 542 GluLeuAlaLysLeuAspPheAsnIleValGlnSerLeuHisGlnGluLeuLysLeu 561	
QY 724 TCCTTCAGGTGTGGAGAATACTGGTGTGTTGTGAGAACTGCCCTTCGACAGGATGA 783	
DB 562 LeuThrArgTrpTrpLysGluSerGlyMetAlaAsp---IleAsnPheThrArgHisArg 580	
QY 784 CTGGTGAATGCTACTTTTGGATATCTGGATCATCGACCCAGCTCGACATCATATTTATGTC 843	
DB 581 ValAlaGluValTyrPhe---SerSerAlaThrPheGluPro---GluTyrSerAlaThr 598	
QY 844 AGATATATGATGGGCAAGTCAACGCTCTGATTCAGTGTGATCGATCATATTTATGTC 903	
DB 599 ArgIleAlaPheThrLysIleGlyCysLeuGlnValLeuPheAspAspMetAlaAspIle 618	
QY 904 TATGGCACCTTAGAAGAACTCGAACCAATTCATCTGACCTCATTCGAGATGGGATTAAC 963	
DB 619 PheAlaThrLeuAspGluLeuLysSerPheThrGluGlyValLysArgTrpAspThrSer 638	
QY 964 TCAATGACCACTTCGCGATACATGCACTGTGCTTCTGCTACTCAACACTTCGTC 1023	
DB 639 LeuLeuHisGluIleProGluCysMetGlnThrCysPheLysValTrpPheLysLeuMet 658	

This SWISS-PROT entry is copyright. It is produced through a collaboration

RESULT 12
TASY TAXCH
ID TASY TAXCH STANDARD; PRT; 862 AA.
AC Q9T37;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Taxadiene synthase (EC 4.2.3.17) (Taxa-4(5),11(12)-diene synthase).
GN TDC1.
OS Taxus chinensis (Chinese yew).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Taxaceae; Taxus.
OX NCBI_TaxID=29808;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Callus;
RA Wang W., Shi Q., Ouyang T., Zhu P., Cheng K.;
RT "Cloning, expression, and characterization of taxadiene synthase, a
RL diterpene cyclase from Taxus chinensis";
CC Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -! FUNCTION: Catalyzes the cyclization of the ubiquitous isoprenoid
CC intermediate geranylgeranyl diphosphate to taxa-4,11-diene, the
CC parent olefin with a taxane skeleton.
CC -! CATALYTIC ACTIVITY: Geranylgeranyl diphosphate = taxa-4,11-diene +
CC diphosphate.
CC -! PATHWAY: Taxol biosynthesis; first step.
CC -! SIMILARITY: Belongs to the terpene synthase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

DE Taxadiene synthase (EC 4.2.3.17) (Taxa-4(5),11(12)-diene synthase).
GN TDC1 OR TASY.
OS Taxis baccata (English yew).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Taxaceae; Taxus.
OX NCBI_TaxID:25629;
RN [1]
RP SEQUENCE FROM N.A.
RA Goerhardt B.;
RL Thesis (2001); Technische Universitaet Berlin, Germany.
CC -!- FUNCTION: Catalyzes the cyclization of the ubiquinoid
intermediate geranylgeranyl diphosphate to taxa-4,11-diene, the
parent olefin with a taxane skeleton.
CC -!- CATALYTIC ACTIVITY: Geranylgeranyl diphosphate = taxa-4,11-diene +
diphosphate.
CC -!- PATHWAY: Taxol biosynthesis; first step.
CC -!- SIMILARITY: Belongs to the terpene synthase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL: AJ320538; CAC42773.1; --
DR InterPro; IPR008930; Terp_cyc_toroid.
DR InterPro; IPR001906; Terp_synth-like.
DR InterPro; IPR005630; Terpene synth C.
DR InterPro; IPR008949; Terpenoid synth.
DR Pfam; PF01397; Terpene synth; 1.
DR Pfam; PF03936; Terpene synth_C; 1.
KW Taxol biosynthesis; Lyase.
FT ACT_SITE 758 835 BY SIMILARITY.
FT ACT_SITE 835 835 BY SIMILARITY.
FT ACT_SITE 839 839 BY SIMILARITY.
SQ SEQUENCE 862 AA; 98048 MW; B37835AFDC1DEFE CRC64;

Alignment Scores:
Pred. No.: 1,43e-38 Length: 862
Score: 634.50 Matches: 164
Percent Similarity: 50.09% Conservative: 109
Best Local Similarity: 30.09% Mismatches: 215
Query Match: 21.82% Indels: 57
DB: 1 Gaps: 18

US-09-938-956-6 (1-1632) x TASY_TAXBA (1-862)
QY 112 ACTTGGTGAAGATGGAAGTGGAGAA-----GAAACG 144
DB 331 ThrLeuLeuAsnLeuLeuLeuAspLysPheGlyGlyCysValProCysMetTyrSerIle 350
QY 145 GATCAAAATTCGACAACTTCAGTTCGATCGATGCTGCAGAGATGGGCTGCCCATCAT 204
DB 351 AspLeuLeuGluLeuSerLeuValAspAsnIleGluHisLeuGlyLeuArgHis 370
QY 205 TTCCAAATAGAGTCAAGAAATCTGTCTCTATATATCTCGACCATCATATTACCAAG 264
DB 371 PheLysGlnGluLeuLysValAlaLeuAspTyrValTyr-----ArgHisTrp----- 386
QY 265 AACCTTTTCCAAAGAGAAAGG-----GAT 291
DB 387 -----SerGluArgGlyIleGlyTrpGlyArgAspSerLeuValProAsp 401
QY 292 CTCTACTCCACATCTCTTCGATTTAGGCTTCCTCAGAGACATGGTTTCAAGTCGCACAA 351
DB 402 LeuAsnThrThraLeuGlyLeuArgThrLeuArgThrHisGlyTyrAspValSerSer 421
QY 352 GAGGTATTCGATAGTTTCAAGAACGAGAGGGGTGAGTTCAAGAAAGCGCTTAGCGAC--- 408
DB 422 AspValLeuAsnAsnPheLysAspGluAsnGlyArgPheSerSerAlaGlyGlnThr 441

QY 409 -----GACACCAGAGGATTGTTGCAACTGTATGAAGCTTCTCTTCTGTTGACGGAAGGC 462
DB 442 HisValGluLeuArgSerValValAsnLeuPheArgAlaSerAspLeuAlaPheProAsp 461
QY 463 GAAACCCAGCTCGAGTCAGCGAGGAATTCCGCCACCAATTTTGGAGGAAAAAGTG---- 519
DB 462 GluGlyAlaMetAspAspAlaArgLysPheAlaGluProTyrLeuArgAspAlaLeuAla 481
QY 520 AACGAGGTGGTGTGATGGCGACCTTTTAAACAAGATCGCATATTCTTTGGACATCCCT 579
DB 482 ThrLysIleSerThrAsnThrLysLeuTyrLysGluLeuGluTyrValGluTyrPro 501
QY 580 CTTCAATTCGAGAGTAATAAGGCCAAATGCACCTGTGTGGATCGAA----- 624
DB 502 TrpHisMetSerIleProArgLeuGluAlaArgSerTyrIleAspSerTyrAspAsp 521
QY 625 -----TGGTATAGGAAG-----AGGCCCGACATG-----AATCCAGTAGTGTG 663
DB 522 TyrValTrpGlnArgLysThrLeuTyrArgMetProSerLeuSerAsnSerLysCysLeu 541
QY 664 GAGCTTCCCATCTCGACTTAAATATTGTTCAAGCACAAATTTCAAGAAGAGCTCAAGAA 723
DB 542 GluLeuAlaLysLeuAspPheAsnIleValGlnSerLeuHisGlnGluLeuLysLeu 561
QY 724 TCCTTCAGGTGGTGGAGAAATACCTGGTGTGTTGTCAGAGCTGCCCTTCGCAAGGGATAGA 783
DB 562 LeuThrArgTrpTrpLysGluSerGlyMetAlaAsp-----IleAsnPheThrArgHisArg 580
QY 784 CTGTGGAATGCTACTTTTGGAAATCTGGGATCATCGAGCCACGTCAGCATGCAAGTGCA 843
DB 581 ValAlaGluValTyrPhe---SerSerAlaThrPheGluPro---GluTyrSerAlaThr 598
QY 844 AGGTAATGATGGCAAGTCAACGCTCTGATTCAGGTGATCGATGATGATATTATGATGTC 903
DB 599 ArgIleAlaPheThrLysIleGlyCysLeuGlnValLeuPheAspAspMetAlaAspIle 618
QY 904 TATGGCACTTAGAAGAACTCGCAACAAATTCACCTGACCTCATTCCGAAGATGGATATAAC 963
DB 619 PheAlaThrLeuAspGluLeuLysSerPheThrGluGlyValLysArgTrpAspThrSer 638
QY 964 TCATCGACCAACTTCCGATTCACGCACTGCTTCTTCTGCACTCAACAACCTCGTC 1023
DB 639 LeuLeuHisGluIleProGluCysMetGlnThrCysPheLysValTrpPheLysLeuMet 658
QY 1024 GATGATACATCTGACGATGTTATGAAGGAGAGGCGTCAACGTTATACCTCTACCTCGCG 1083
DB 659 GluGluValAsnAsnAspValValLysValGlnGlyArgAspMetLeuAlaHisIleArg 678
QY 1084 CAATCGTGGTGGATTGGCGGATAGTATATGGTAGAGCAGCGGTGCTCTACGCGGG 1143
DB 679 LysProTrpGluLeuTyrPheAsnCysTyrValGlnGluArgGluTrpLeuGluAlaGly 698
QY 1144 CACAAACCAAGTTTGAAGAGATTTGGAGAACTCATGCGCAGTCGATAAGT---GGGCCCC 1200
DB 699 TyrIleProThrPheGluGluTyrLeuLysThrTyrAlaIleSerValGlyLeuGlyPro 718
QY 1201 TGTATGTTAAACGCACATATTCTTCCGAGTAGAACATTCGTTCTCAAGAGACCGTCGAC 1260
DB 719 CysThrLeuGlnProIleLeuLeu---MetGlyGluLeuValLysAspAspValValGlu 737
QY 1261 AGTTTGTACAAATACCAGATTTAGTTCGTGTGTCATCTCTCTGCTGCGGCTTGCTGAT 1320
DB 738 LysValHisTyrProSerAsnMetPheGluLeuValSerLeuSerTrpArgLeuThrAsn 757
QY 1321 GATTTCGGAACCTCGGTGGAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGC 1380
DB 758 AspThrLysThrTyrGlnAlaGluLysAlaArgGlyGlnGlnAlaSerGlyIleAlaCys 777
QY 1381 TACATGAGTCACTAC---AATGATCGAGCGCGGCGGAGGCGGAGGCGTGAATGCTG 1437
DB 778 TyrMetLysAspAsnProGlyAlaThrGluLysAlaIleLysHisIleCysArgVal 797
QY 1438 ATACCGGAGGTGTCGAAGAGATGATTCGCGAG-----AGGGTGTGCGAGGATTCCTCCA 1491

```

Db 798 ValAspArgAlaLeuLysGluAlaSerPheGluTyrPhelysProSerAsnAspIlePro 817
QY 1492 TTCGGC---AAAGATTTTATAGGATGTCAGTGTGATTTAGGAAGGATGGCGAGTTGATG 1548
Db 818 MetGlyCysLysSerPheIle-----PheAsnLeuArgLeuCysValGlnIlePhe 834
QY 1549 TACCATAATGAGATGGCGGCGACACACACACACCTATTATACATCAACAAATGACGAGA 1608
Db 835 TyrLysPheIleAspGlyTyrGlyIleAlaAsnGluIleLysAspTyrIleArgLys 854
QY 1609 ACCTTATTCGAGCC 1623
Db 855 ValTyrIleAspPro 859

RESULT 14
YK82_YEAST
ID YK82_YEAST STANDARD; PRT; 1169 AA.
AC P36170;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical 122.2 kDa protein in SIR1 3' region precursor.
GN YK102W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Gallion L., Dujon B.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Cell wall (By similarity).
CC -!- PTM: Extensively O-glycosylated (By similarity).
CC -!- SIMILARITY: STRONG, TO YEAST PROTEIN FLO1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z28327; CAA8182.1; -.
CC DR PIR; S38181; S38181.
CC DR GeneOnline; 140081; -.
CC DR SGD; S0001810; FLO10.
CC KW Hypothetical protein; Glycoprotein; Repeat; Cell wall; Signal.
CC FT SIGNAL 1 26
CC FT CHAIN 27 1169
CC FT DOMAIN 274 609
CC FT CARBOHYD 122 122
CC FT CARBOHYD 157 157
CC FT CARBOHYD 279 279
CC FT CARBOHYD 389 389
CC FT CARBOHYD 452 452
CC FT CARBOHYD 515 515
CC FT CARBOHYD 578 578
CC FT CARBOHYD 656 656
CC FT CARBOHYD 686 686
CC FT CARBOHYD 879 879
CC FT CARBOHYD 1092 1092
CC FT CARBOHYD 1099 1099
CC FT CARBOHYD 1169 AA; 122164 MW; 940C6DFB3569C669 CRC64;
CC SQ SEQUENCE

Alignment Scores:
Pred. No.: 0.0232 Length: 1169
Score: 138.50 Matches: 105
Percent Similarity: 39.29% Conservative: 73
Best Local Similarity: 23.18% Mismatches: 164
Query Match: 1 Indels: 112
DB: 21 Gaps: 21

```

```

US-09-938-956-6 (1-1632) x YK82_YEAST (1-1169)
QY 1558 CATTATGTCATCACTCACTCGCG-----CCATCCTTCTCTAAATCAACTGCAC 1514
Db 569 TyrValSerSerSerThrAlaAlaAsnTyrThrSerSerPheSerSerSerGlu 588
QY 1513 ATCCTATAAATCTTTCGCGAATGAGAAATCTTCGACACCCCTCTCCGCAATCATCTCT 1454
Db 599 ValCysThrGluCysThrGluThrSerThrSerThrProTyrProTyrAlaThrSer 608
QY 1453 TCCACACC-TCCGCTATCAGCATTTCAGCTTCGCGCTCCGCTCCGCTCCGATGCAATG 1395
Db 609 SerThrGlyThrAla-----ThrSerPheThrAlaSerThrSerAsnThrMet 624
QY 1394 TAGTCATCATGTAGTACACTGAAGT-----GATTTCGGCACATCCCTCTCTCTCACTCT 1341
Db 625 ThrSerLeuValGlnThrAspThrValSerPheSerLeuSerThrValSerGlu 644
QY 1340 TCCACCGAGGTTCCCAATCATCATCAGCAGCGCAGA-----ACGAAGGAT 1296
Db 645 HisThrAsnAlaProThrSerSerValGluSerAsnAlaSerThrPheIleSerSerAsn 664
QY 1295 GACCAACGAACATAATCGTGTATTGTACAAACTGTCAGCGTCTCTCTTTGTG----- 1242
Db 665 LysGlySerValLysSerTyrValThrSerSerIleHisSerIleThrProMetTyrPro 684
QY 1241 ---AACGAATCTGTTACTCGGAAGATATGCGTTAAACATACAGAGGCCCACTTATCGAC 1185
Db 685 SerAsnGlnThrValThrSerSerVal-----ValSerThrProIleThrSer 701
QY 1184 TCCCATGAGTTCTCCAAATATCTTCCAAATCTTGTGTTGTCGCGCGTAGACACCGCT 1125
Db 702 -----GluSerSerGluSerSerAlaSerValThrIleLeuPro----- 714
QY 1124 GCCTTACCATATCTATTCGCCCAA---TCAACCCACGATTGCGCGAGTAGGGTATA 1068
Db 715 ---SerThrIleThrSerGluPheLysProSerThrMetLysThrLysValValSerIle 733
QY 1067 ACGTTGACGCTCTCTCTTCAATACATCGTACCATGTATCATCGACGAGTTGTTGAGT 1008
Db 734 SerSerSerProThrAsnLeuIleThrSerTyrAspThrThrSer----- 748
QY 1007 CCAAGAAAGCAGCTGTCATGCTAATCGGAAAGTTGGTGGATGAGTTTATATCCACTT 948
Db 749 ---LysAspSerThrValGlySerSerThrSerValSerLeuIleSerSerIle 766
QY 947 CGAATGAGTTCAGTGAATTGTTTCGAGTTCTTCTAAGGTGCCATAGACATCAATAATCA 888
Db 767 SerLeuProSerSerTyrSerAlaSerSerGluGlnIlePheHisSerSerIleValSer 786
QY 887 TCGATCACCGTAATCAGAGCGTTGACTTTGCGCATCATATTCCTTGCACTTGCAATGCTGA 828
Db 787 Ser----- 787
QY 827 CGTGCCTCGATGCCAGTATTCCAAAGTAGCATCCACAGTCTATCCCTTCCGGAAG 768
Db 788 AsnGlyGlnAlaLeuThrSerPheSer-----SerThrLysValSerSerSerGlu 804
QY 767 GGCAGCTTCTCAACAAACCCAGTATTTCTCCACCACCTGGAAGGATTTCTTGAGCTTCT 708
Db 805 SerSerGluSer-----HisArgThrSerProThrThrSerSerGlu 818
QY 707 TGAATTTGTGCTGACAAATATTTAAGTCGAGTATGCGACAGTCCACACACTACTGATT 648
Db 819 -----SerglyIleLysSerSerGlyValGluIleGluSerThrSerThr 833
QY 647 ATGTCGGCGCTCTTCTCTATACCATTCGATCCACACAGGTGCATTTGGCTTTTAACTCTC 588
Db 834 SerSer-----PheSerPheHisGluThrSerThrAlaSerThrSerValGlnIle--- 850
QY 587 CAATGAAGAGGATGTCCAAAGAATATGCGATTCTTGTATAAAGTCGCCATGCACACCA 528

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 25, 2004, 05:37:34 ; Search time 79.5 Seconds
(without alignment)
11600.441 Million cell updates/sec

Title: US-09-938-956-6

Perfect score: 2908

Sequence: 1 atgagagatccggaacta.....tattcgagccctttgcatga 1632

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 3172214

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-O=/cgn2.1/USPTO.epool.p/US0938956/runat_24022004.144659.3047/app.query.fasta_1.1799
-DB=A_Geneseq_29Jan04 -OPMT=fastan -SUFFIX=rag -MINMATCH=0 -1-LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pcp -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0938956 @CGN 1.154 @runat_24022004.144659.3047 -NCFU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_29Jan04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001s.*
- 5: Geneseq2002s.*
- 6: Geneseq2003s.*
- 7: Geneseq2003bs.*
- 8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2881	99.1	543	ABB75749	Mentha sp
2	2876	98.9	599	AAR74291	Aar74291 Spearmint
3	2876	98.9	599	AAy90838	AAy90838 Spearmint
4	2876	98.9	599	AAy70716	AAy70716 Mint 4S-1
5	2876	98.9	599	AAE29781	AAE29781 M. spiccat
6	2825	97.1	599	AAE29400	AAE29400 Spearmint
7	1636.5	56.3	591	AAy02332	AAy02332 Computer
8	1634.5	56.2	591	AAy02335	AAy02335 Computer
9	1633.5	56.2	591	AAy02309	AAy02309 A sage 1,
10	1633.5	56.2	591	AAy02331	AAy02331 Computer

11	1633.5	56.2	591	3	AAy90839	Sage 1,8
12	1632.5	56.1	591	2	AAy02336	Computer
13	1632.5	56.1	591	2	AAy02329	Computer
14	1631.5	56.1	591	2	AAy02330	Computer
15	1631.5	56.1	591	2	AAy02337	Computer
16	1631.5	56.1	591	2	AAy02333	Computer
17	1630.5	56.1	591	2	AAy02338	Computer
18	1630.5	56.1	591	2	AAy02334	Computer
19	1553.5	53.4	589	2	AAy02310	A sage (+
20	1553.5	53.4	590	6	AAE90857	Sage +sac
21	1553.5	53.4	590	6	AAE29782	S. offic
22	1552.5	53.4	590	2	AAy02321	Computer
23	1552.5	53.4	590	2	AAy02322	Computer
24	1551.5	53.4	590	2	AAy02325	Computer
25	1551.5	53.4	590	2	AAy02323	Computer
26	1550.5	53.3	590	2	AAy02326	Computer
27	1550.5	53.3	590	2	AAy02328	Computer
28	1550.5	53.3	590	2	AAy02324	Computer
29	1550.5	53.3	590	2	AAy02327	Computer
30	1547.5	53.2	589	2	AAy02320	Computer
31	1471	50.6	598	2	AAy02314	Computer
32	1469	50.5	598	2	AAy02317	Computer
33	1469	50.5	598	2	AAy02318	Computer
34	1468	50.5	598	2	AAy02308	Computer
35	1468	50.5	598	3	AAy90840	Sage born
36	1467	50.4	598	2	AAy02316	Computer
37	1467	50.4	598	2	AAy02311	Computer
38	1466	50.4	598	2	AAy02312	Computer
39	1466	50.4	598	2	AAy02319	Computer
40	1466	50.4	598	2	AAy02313	Computer
41	1465	50.4	598	2	AAy02315	Computer
42	1203.5	41.4	583	5	AAm50097	Armopha-4
43	1161.5	39.9	562	3	AAg51896	Arabidops
44	1161.5	39.9	566	3	AAg51895	Arabidops
45	1161.5	39.9	598	3	AAg51894	Arabidops

ALIGNMENTS

RESULT 1

ABB75749

ID ABB75749 standard; protein; 543 AA.

AC ABB75749;

XX XX

DT 24-JUN-2002 (first entry)

XX XX

DE Mentha spicata limonene synthase.

XX XX

KW Limonene synthase; enzyme; monoterpene; cyclic terpenoid; flavour; fragrance; plant.

XX XX

OS Mentha spicata.

XX XX

PN WO200220815-A2.

XX XX

PD 14-MAR-2002.

XX XX

PF 28-AUG-2001; 2001WO-US026853.

XX XX

PR 01-SEP-2000; 2000US-0229858P.

XX XX

PR 01-SEP-2000; 2000US-0229907P.

XX XX

PA (DUPO) DU PONT DE NEMOURS & CO E I.

XX XX

PI Dicosimo DJ, Koffas M, Wang S;

XX XX

DR WPI; 2002-339805/37.

XX XX

DR N-PSDB; ABL53856.

XX XX

PT Producing cyclic terpenoids, e.g. monoterpenes, which are useful in the

PT fragrance or pharmaceutical industry, by employing a transformed

PT methanotrophic bacterium that metabolize single carbon substrates as a

XX AAR74291;
 XX
 XX 25-MAR-2003 (revised)
 XX 24-OCT-1995 (first entry)
 XX
 XX Spearmint limonene synthase.
 XX
 XX LS; expression; production; predator defence mechanism; enzyme.
 XX OS
 XX Mentha spicata.
 XX OS
 XX WO9511913-A1.
 XX PN
 XX 04-MAY-1995.
 XX PD
 XX 27-OCT-1994; 94WO-US012379.
 XX PF
 XX 28-OCT-1993; 93US-00145941.
 XX PR
 XX
 XX (UNIW) UNIV WASHINGTON STATE RES FOUND.
 XX PA
 XX Croteau RB, Colby SM;
 XX X
 XX WPI; 1995-178818/23.
 XX DR
 XX N-PSDB; AAQ90066.
 XX
 XX Nucleotide sequence for spearmint limonene synthase - useful as predator
 XX PT defence mechanism for plants.
 XX PR
 XX Claim 7; Fig 4; 62pp; English.
 XX
 XX The sequence is that of the spearmint limonene synthase protein. The
 XX CC enzyme can be used in plants as a predator defence mechanism. (Updated on
 XX CC 25-MAR-2003 to correct PN field.)
 XX
 XX Sequence 599 AA;
 XX SQ

Alignment Scores:
 Pred. No.: 3,566-266 Length: 599
 Score: 2876.00 Matches: 542
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98,90% Indels: 0
 DB: 2 Gaps: 0

US-09-938-956-6 (1-1632) x AAR74291 (1-599)

QY 4 AGACGATCGGAACTACAAACCTTCTCGTTGGGATGTCAACTTCATCCAAATCGCTTCTC 63
 DB 58 ArgArgSerGlyAsnTyrAsnProSerArgTrpAspValAsnPhelieGlnSerLeuLeu 77
 QY 64 AGTGACTATAAGGAGGACAAACACGTGATTAGGCTCTGAGCTGGTCACTTTGGTGAAG 123
 DB 78 SerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuValLys 97
 QY 124 ATGGAACCTGGAGAAAGAAACGGATCAAAATTCGACAACTTGAGTTGATCGATGATTCGAG 183
 DB 98 MetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspAspLeuGln 117
 QY 184 AGGATGGGGCTCGCGATCATTCCAAATAGTTTCAAGAAATCTGCTCTATATAT 243
 DB 118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPhelLysGluIleLeuSerSerIleTyr 137
 QY 244 CTCGACCATCATATTACAGAACCTTTTCCAAAAGAAAGAGGATCTCTACTCCACA 303
 DB 138 LeuAspHisIleTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157
 QY 304 TCTCTGTGCTTTAGGCTCTCAGAGAACATGGTTTCAAGTCGACACAGAGGATTCGAT 363
 DB 158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177
 QY 364 AGTTTCAAGAACGAGGAGGTGAGTTCAAGAAAGCCTTAGCCGACGACACGAGGATTG 423

DB 178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197
 QY 424 TTGCAACTGTATGAAGCTTCTTTCTGTGTGAGGAGCGGAAACACGCTCGAGTCAGCG 483
 DB 198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217
 QY 484 AGGGAATTGCGCCACCAAAATTTTGGAGGAAAAAGTGAACGAGGGGTGGTGTGTGATGGCGAC 543
 DB 218 ArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyGlyValAspGlyAsp 237
 QY 544 CTTTAAACAAGAAATCGCATATCTTTGGACATCCCTCTTCAATGGAGGATTAAGAGCCCA 603
 DB 238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro 257
 QY 604 AATGCACCTGTGTGATCGAATGTATAGGAAGAGGCCCGACATCAATCCAGTAGTGTG 663
 DB 258 AsnAlaProValTrpIleGluTrpTyrArgLysArgProAspMetAsnProValValLeu 277
 QY 664 GAGCTTGGCCTACTCGACTTAAATATCTTCAACACAAATTTCAAGAGAGCTCAAGAA 723
 DB 278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluGluLeuLysGlu 297
 QY 724 TCCTTCAGGTGGTGGAGAAATACTGGGTGTGTGTGAGAAAGCTGCCCTTCGCAAGGATAGA 783
 DB 298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317
 QY 784 CTGGTGGATGCTACTTTTGAATCTCGGATCATCGAGCCACGCTCAGCATCGAAGTGA 843
 DB 318 LeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla 337
 QY 844 AGATAATGATGGGCAAAAGTCAACGCTCTGTATTAAGGTGATCGATGATATTTATGATGC 903
 DB 338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357
 QY 904 TATGGCACCTTAGAAGAACTCGAACTCACTCACTCACTCACTCACTCACTCACTCACTCA 963
 DB 358 TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgTrpAspIleAsn 377
 QY 964 TCAATCGACCAACTTCCGATTAATGATCAACTGTGCTTCTTTCGACATCAACAACTTCGTC 1023
 DB 378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPhaVal 397
 QY 1024 GATGATACATCGTACGATGTTATGAAGGAGAAAGCGCTCAACGTATACCTACCTACCTCGG 1083
 DB 398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417
 QY 1084 CAATCGTGGTGTGATTTGGCGGATAAGTATATATGATAGAGCACGCTGCTTCTACGGCGGG 1143
 DB 418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437
 QY 1144 CACAAACCAAGTTTGGAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCGCTGT 1203
 DB 438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457
 QY 1204 ATGTTAAACGACATATTTCCGAGTAAACAGATTCTGTTCAAAAGGAGACCGTCGACAGT 1263
 DB 458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477
 QY 1264 TTGTACAAATACCAAGTATTTAGTTTGTGTGCTTCCTTCCTTCGCTGCTGCTGATGAT 1323
 DB 478 LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497
 QY 1324 TTGGAACTCGGTGGAGAGGTGAGCAGAGGGGATGTGCCGAAATCACTTCAGTGTCTAC 1383
 DB 498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517
 QY 1384 ATGAGTGACTACATGATCGAGCGGCGGAGCGGAGCAGCAGTGAATCGCTGATGATGCG 1443
 DB 518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValIleTyrPheIleAla 537
 QY 1444 GAGGTGTGAAGAAAGATGAATTCGCGAGAGGGGTGTGGAAGGATTCCTCATTCGCAAGAT 1503

538 GluValTrpLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557
 1504 TTTATAGATGTCGATTCATTTAGGAGGATGGCGCAGTTGATGTACCATATGAGAT 1563
 558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577
 1564 GGGCAGCGCAGCACACACCTATTATACATCAACAATGACAGAACCTTATTCGAGGCC 1623
 578 GlyHisGlyThrGlnHisProIleHisGlnGlnMetThrArgThrLeuPheGluPro 597
 1624 TTTGCA 1629
 598 PheAla 599
 RESULT 3
 AAY90838
 ID AAY90838 standard; protein; 599 AA.
 AC AAY90838;
 XX
 25-AUG-2000 (first entry)
 XX
 DE Spearmint limonene synthase protein sequence SEQ ID NO:22.
 XX
 KW Synthese; protein co-ordinate data; active site; modification; terpenoid;
 KW 3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;
 KW isoprenoid; breeding programme; fragrance; flavour; pheromone;
 KW defensive agent; pigment; antitumor; steroid hormone;
 KW signal transduction pathway; bile acid; affinity purification;
 KW photoreceptor; enzymatic synthesis; nutrient supplement;
 KW immunological reagent.
 XX
 OS Mentha spicata.
 XX
 PN WO200017327-A2.
 XX
 PD 30-MAR-2000.
 XX
 PF 17-SEP-1999; 99WO-US021419.
 XX
 PR 18-SEP-1998; 98US-0100993P.
 PR 22-APR-1999; 99US-0130628P.
 PR 23-AUG-1999; 99US-0150262P.
 XX
 XX (KENT) UNIV KENTUCKY RES DEPT.
 PA (SALK) SALK INST BIOLOGICAL STUDIES.
 PA
 XX Chappell J, Manna KR, Noel JP, Starks CM;
 PI
 XX WPI; 2000-292839/25.
 DR N-PSDB; AAA38923.
 XX
 XX Novel terpene synthase enzymes, useful for producing terpene
 PT hydrocarbons, e.g. fragrances or antitumor agents, are derived from known
 PT enzymes by specific amino acid alterations.
 PT
 XX Claim 103; Page 371-373; 450pp; English.
 XX
 XX The present invention describes an isolated terpene synthase (I)
 CC comprising a region with at least 20% identity to region 265-535 of a 548
 CC amino acid (aa) sequence (Ia), given in AAY90831. (I) contains nine alpha
 CC -carbon atoms (alphaC) that have interatomic distances, between each
 CC other, within tabulated ranges, have a centre point (within a sphere of
 CC radius 2.3 Angstrom) within tabulated ranges, and have an ordered
 CC arrangement of R groups (defining aa side chains), excluding specific
 CC tabulated arrangements (tables given in the specification). (I), and
 CC related enzymes, are used to produce a wide range of terpenoids (e.g.
 CC cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances,
 CC flavours, pheromones, defensive agents, pigments, antitumor agents,
 CC components of signal transduction pathways, precursors of steroid
 CC hormones and bile acids, as photoreceptors and as co-factor side chains.
 CC Some synthases with little or no catalytic activity (and nucleic acids
 CC encoding them) are used as controls in the analysis of products formed by

CC enzymatic synthesis; as nutrient supplements; for affinity purification
 CC of isoprenoids; or to develop immunological reagents or nucleic acids for
 CC monitoring expression of terpene synthase or inheritance of the gene in
 CC plant breeding programs. The new synthases may produce novel terpene
 CC products. AAA38910 to AAA38938 and AAY90831 to AAY90859 represent
 CC sequences used in the exemplification of the present invention
 XX
 SQ Sequence 599 AA;
 Alignment Scores:
 Pred. No.: 3,56e-266 Length: 599
 Score: 2876.00 Matches: 542
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.90% Indels: 0
 DB: 3 Gaps: 0
 US-09-938-956-6 (1-1632) x AAY90838 (1-599)
 QY 4 AGACGATCCGAAACTACAAACCTTCTGTTGGATGTCAACTTCATCCATCGCTTC 63
 DB 58 ArgArgSerGlyAsnTyrAsnProSerArgTrpAspValAsnPheIleGlnSerLeuLeu 77
 QY 64 AGTGACTATAGGAGGACAAACACGTCATTAGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
 DB 78 SerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuValLys 97
 QY 124 ATGGAACCTGAGAAAGAAACCGATCAAAATTCGACAACTTGAGTTGATCGATGCTCCAG 183
 DB 98 MetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspAspLeuGln 117
 QY 184 AGGATGGGCTGTCGGATCATTTCCAAAATGAGTTCAAAAGAAATCTGTCTCTATATAT 243
 DB 118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluLeuLeuSerSerIleTyr 137
 QY 244 CTCGACCATCACTATTACAAAGAACCCCTTTCCAAAAGAAAGAGGATCTCTACTCCACA 303
 DB 138 LeuAspHisHisTyrTyrLysAsnProPheProLysGluArgAspLeuTyrSerThr 157
 QY 304 TCTCTTGATTTAGCTCCTCAGAGAACATGGTTTCAAGTCCGACAGAGGATTCGAT 363
 DB 158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177
 QY 364 AGTTTCAAGACGAGGAGGAGTTCAGAAAGACCTTAGCGACGACACACAGAGGATTG 423
 DB 178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197
 QY 424 TTGCAACTGTATGAAGCTTCTCTGTTGACGAGGCGAAACACGCTCGACTCAGCG 483
 DB 198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217
 QY 484 AGGGAATTCGCCACCAAAATTTTTCGAGGAAAAGTGAACGAGGCTGTGTGATGGCGAC 543
 DB 218 ArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyValAspGlyAsp 237
 QY 544 CTTTAAACAAGAATCGCATATTCTTTGGACATCCCTCTTCAATTCGAGGATTAAAGGCCA 603
 DB 238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro 257
 QY 604 AATGCACTGTGCGATCGATGCTATAGGAGAGGCCCGACATGAATCCAGTAGTGTG 663
 DB 258 AsnAlaProValTrpIleGluTrpTyrArgLysArgProAspMetAsnProValValLeu 277
 QY 664 GAGCTTGCCATATCGATTCGATTAAATATTGTTCAAGCACAAATTTCAAGAGAGCTCAAAGAA 723
 DB 278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluLeuLysGlu 297
 QY 724 TCTTCAGGTGGTGAGAAATACATCGGGTTTGTGAGAGCTGCCCTTCGAGGATAGA 783
 DB 298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317
 QY 784 CTGTGGGAATGCTACTTTTGGAACTCTGGATCATCGAGCCACGCTCAGCATCGCAAGTGCA 843

QY 424 TTGCAACTGTATGAAGTCTCTTCTGTTGACGGAGGCGAAACACACGCTCGAGTCAGCG 483
Db 198 LeuGlnLeuTyrGlnAlaSerPheLeuLeuThrGluGlyGluThrLeuLeuSerAla 217
QY 484 AGGGAATTCGCCACCAAAATTTTGGAGGAAAGTGAACGAGGCTGGTGTTCATCGGCAC 543
Db 218 ArgGluPheAlaThrLysPheLeuGluLysValAsnGluGlyValAspGlyAsp 237
QY 544 CTTTAAACAAGATCCATATCTTTGGACATCCCTCTTCATTCGAGGATTAAGAGCCA 603
Db 238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro 257
QY 604 AATGCACCTGTGTGGATCGAATGTATAGGAAGCGCCGACATGAATCCAGTACTGTGTG 663
Db 258 AsnAlaProValTrpIleGluTrpTyrArgLysArgProAspMetAsnProValValLeu 277
QY 664 GAGCTTGCCATCTCGCACTTAATATTGTTCAGACAAATTTCAAGAGAGCTCAAGAA 723
Db 278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluLysGlu 297
QY 724 TCCTTCAGGTGGGAGAAATACTGGTGTGTGTGAGAGCTGCCCTCGCAGGATAGA 783
Db 298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317
QY 784 CTGGTGGAAATGCTACTTTTGGAACTATCGGATCATCGACCCAGCTCAGCATGCAAGTCA 843
Db 318 LeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla 337
QY 844 AGGATATGATGGCAAGTCAACGCTCTGATTACGGTATCGATGATATTTATGATGC 903
Db 338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357
QY 904 TATGGCACCTTAGAAGCACTCGAAACAATTCACCTGACCTCATTCGAAGATGGGATATAAC 963
Db 358 TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgArgTrpAspIleAsn 377
QY 964 TCAATGCACCAATTCGGATTTACATGCAACTGTGCTTCTTGGCACTCAACAATTCGTC 1023
Db 378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397
QY 1024 GATGATACATCGTACCATGTTATGAAGGAGAAAGCGTCAACGTTATACCTTACCTGGG 1083
Db 398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417
QY 1084 CAATCTGGGTGATTTGGCGAATAAGTATATGTTAGAGGACCGTGTCTTACCGCGGG 1143
Db 418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437
QY 1144 CACAAACCAAGTTTGAAGAGTATTTGGAAGACTCATGCGACGATAGTGGCGCCTGT 1203
Db 438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457
QY 1204 ATGTTAACGCACATATCTTCCGAGTAAACAGATTGTTCCAAAGAGACCGTCGACAGT 1263
Db 458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477
QY 1264 TTGTACAAATACCAATATTAGTTGTTGTTGTCATCTCTGTTCTCGGCTTCTGATGAT 1323
Db 478 LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497
QY 1324 TTGGGAACCTCGGTGAAAGAGTGTGACAGAGGAGTGTGCCGAATACCTTCAGTGTAC 1383
Db 498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517
QY 1384 ATGAGTGTACATATCGATCGAGGCGGAGCGCGGAGCAGCTGAATGGCTGTATACGG 1443
Db 518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValTyrLysIleAla 537
QY 1444 GAGGTGTGAAGAAGATGAATCGGAGAGGCTGTGCAAGGATTCCTCCATTCGGCAAGAT 1503
Db 538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557
QY 1504 TTTATAGGATGTCAGTTGATTTAGGAAGGATGGCGCAGTTGATGATCAATTAATGAGAT 1563

Db 558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577
QY 1564 GGGCACGCGCACAAACACACCCCTATTATACATCAACAATGACCAAGACCTTATTTCGAGCCC 1623
Db 578 GlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGluPro 597
QY 1624 TTTCGA 1629
Db 598 PheAla 599
RESULT 5
AAE29781
ID AAE29781 standard; protein; 599 AA.
XX AAE29781;
AC AAE29781;
XX 24-FEB-2003 (first entry)
DT XX
XX M. spicata limonene synthase.
DE XX
XX Insect-assisted cross-pollination; scent biosynthetic enzyme; enzyme;
KW hybrid seed; limonene synthase.
XX
XX Mentha spicata.
OS
XX WO200276189-A1.
PN
XX 03-OCT-2002.
PD
XX 24-FEB-2002; 2002WO-IL000142.
PF
XX 22-MAR-2001; 2001US-0277657P.
PR
XX (SCEN-) SCENTGENE POLLINATION LTD.
PA
XX
XX Paldi N;
PI
XX WPI; 2003-018835/01.
DR N-PSDB; AAD47174.
XX
XX Enhancing insect assisted cross-pollination between flowering plants
PT useful for producing hybrid seeds, comprises co-expressing at least one
PT scent biosynthetic enzyme in plants and growing them in a vicinity with
PT pollinating insect.
XX
XX Disclosure; Page 70-71; 89pp; English.
XX
XX The invention relates to a method of enhancing insect-assisted cross-
CC pollination between flowering plants of a single plant species that have
CC at least 2 different genetic backgrounds (e.g. different cultivars). The
CC method involves co-expressing in plants at least one scent biosynthetic
CC enzyme and growing the plants in a cross-pollination vicinity in the
CC presence of at least one pollinating insect. The method is useful for the
CC production of hybrid seeds. The present sequence is M. spicata limonene
CC synthase, a scent biosynthetic enzyme
XX
SQ Sequence 599 AA;
Alignment Scores:
Pred. No.: 3,56e-266 Length: 599
Score: 2876.00 Matches: 542
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.90% Indels: 0
DB: 6 Gaps: 0
US-09-938-956-6 (1-1632) x AAE29781 (1-599)
QY 4 AGAGATCCGGAACACTACACCTTCTCGTTGGATGTCACTTCATCCAACTCGTCTTC 63
Db 58 ArgArgSerGlyAsnTyrAsnProSerArgTrpAspValAsnPheIleGlnSerLeu 77

QY 64 AGTGACTATAAGGAGGACAAACACGCTGATTAGGGCTTCGAGCTGGTCACCTTGGTGAAG 123
Db 78 SerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuValLys 97
QY 124 ATGGAACCTGAGAAAGAAACGGATCAAAATTCACAACTTGAATTCATCGATCGATCGAG 183
Db 98 MetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspLeuGln 117
QY 184 AGGATGGGCTGCTCGATCATTTCCAAATAGATTCAAAGAAATCTTCTCTATATAT 243
Db 118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIleTyr 137
QY 244 CTCGACCATCACTATTACAAAGAACCTTTTCCAAAGAAAGAGGATCTCTACTCCACA 303
Db 138 LeuAspHisIleTyrTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157
QY 304 TCTCTTGATTTAGGCTCTCCAGAGAACATGTTTTCAAGTCGACAAAGAGATATCGAT 363
Db 158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177
QY 364 AGTTTCAAGAACGAGGAGGTGAGTTCAAGAAAGCCTTAGCGACGACACAGAGGATTG 423
Db 178 SerPheLysAsnGluGluGluPheLysGluSerLeuSerAspThrArgGlyLeu 197
QY 424 TTGCAACTGATCAAGCTTCTCTTCTGTTGACGAGGCGAAACACACGCTCGAGTCAGCG 483
Db 198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGlnSerAla 217
QY 484 AGGGAATTCGCCACCAATTTTGAGGAAAGTGAACGAGGGTGTGTGATGGCGAC 543
Db 218 ArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyGlyValAspGlyAsp 237
QY 544 CTTTAAACAAGAAATCGCATATCTTTGGACATCCCTCTTCTATGAGGATTAAGGCGCA 603
Db 238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro 257
QY 604 ATGACCTGTGTGATCGAATGTATAGGAAGCGCCGACATCAATCCAGTGTGTG 663
Db 258 AsnAlaProValTrpIleGluTrpTyrArgLysArgProAspMetAsnProValLeu 277
QY 664 GAGCTTGCCATCTCGCTTAATATTTCAAGCACAAATTCAGAGAGAGCTCAAGAA 723
Db 278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluLeuLysGlu 297
QY 724 TCTTTCAGTGTGGAGAAATCTGGTTTGTGTGAGAAGCTGCCCTTCGCAAGGATAGA 783
Db 298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317
QY 784 CTGGTGGATGTACTTTGGATCTGGATCTCGACGACCGCCAGCTCAGCATGCCAGTCCA 843
Db 318 LeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla 337
QY 844 AGGATAATGATGGCAAGCTCAACGCTCTGATTACGCTGATCGATGATATTATGATGC 903
Db 338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357
QY 904 TATGGCACCTTGAAGAACTCGAACATTCATCTGACCTCATTCGAAAGATGGATATAAC 963
Db 358 TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgArgTrpAspIleAsn 377
QY 964 TCATTCGACCAACTCCCGATTAATGCAACTGTCTTCTTGTGCACTCAACAACTTCGTC 1023
Db 378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397
QY 1024 GATGATACATCTGATGATTTAGGAGGAGAAAGCGTCAACGTTATACCTTACCTGCGG 1083
Db 398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417
QY 1084 CAATCGTGGTGTGATTCGCGATTAAGTATATGTAGAGGACGCGTGTCTTACGGCGGG 1143
Db 418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437
QY 1144 CACAACCAAGTTTGAAGAGATTTTGGAGAACTCATGCGACATCATGAGTGGGCGCTGT 1203

Db 438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457
QY 1204 ATGTTTAACGACACATATCTTCCGAGTAACACAGATTCGTTCAAAAGGAGACCGTCGACAGT 1263
Db 458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477
QY 1264 TTGTACAAATACACAGATTTAGTTCGTGTGTCATCCCTTCGTTCTGGCGCTTGTCTGATGAT 1323
Db 478 LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497
QY 1324 TTGGAACTTCGCTGGAAGAGGTGAGCAGAGGGGTGTCGCGAAATCACATTCAGTGTCTAC 1383
Db 498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517
QY 1384 ATGAGTGACTACAAATGTCATCGAGCGGAGGCGCGAAGCACGTAATAATGGCTGATAGCG 1443
Db 518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla 537
QY 1444 GAGGTGTGGAAGAAGATGATGCGGAGAGGGTGTGCAAGGATTCCTCATTCGCGCAAAGAT 1503
Db 538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557
QY 1504 TTTATAGGATGTGAGTTGATTTAGGAAGGATGCGGAGTGTGATGATACCATATGAGAT 1563
Db 558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577
QY 1564 GGGCAGGCGACACAAACCCCTATTATACATCACAAATGACACAGAACCTTATTTCGAGCCC 1623
Db 578 GlyHisGlyThrGlnHisProIleIleHisGlnMetThrArgThrLeuPheGluPro 597
QY 1624 TTTGCCA 1629
Db 598 PheAla 599
RESULT 6
AAB29400
ID AAB29400 standard; protein; 599 AA.
AC AAB29400;
XX XX
DT 09-FEB-2001 (first entry)
XX XX
DE Spearmint limonene synthase.
XX XX
KW Spearmint limonene synthase; transgenic microorganism;
KW monoterpene production; geranyl diphosphate synthase activity;
KW monoterpene synthase.
OS Mentha spicata.
XX XX
FN JP2000245482-A.
XX XX
PD 12-SEP-2000.
XX XX
PF 05-MAR-1999; 99JP-00059431.
XX XX
PR 05-MAR-1999; 99JP-00059431.
XX XX
PA (SOZO-) SOZOTEKI SEIBUTSU KOGAKU KENKYUSHO KK.
XX XX
WPI; 2000-642026/62.
DR N-PSDB; AAC63810.
XX XX
XX New transformed microbe useful for the production of monoterpene.
XX Example; Page 10-13; 15pp; Japanese.
CC The invention relates to a transformed microorganism containing a first
CC gene encoding a protein with geranyl diphosphate synthase activity, and a
CC second gene encoding a monoterpene synthase. The microorganism is able to
CC produce the particular monoterpene whose synthesis is catalysed by the
CC monoterpene synthase that has been introduced. The invention also relates

CC to the production of monoterpenes using the novel microorganism. The
CC microorganism can be used for the production of various monoterpenes
CC which are useful in industry. In an exemplification of the invention,
CC *Escherichia coli* DH5-alpha was transformed with cDNA encoding spearmint
CC limonene synthase (AAC63810) and DNA encoding *Bacillus stearothermophilus*
CC farnesyl diphosphate synthase (AAC63809). The present sequence represents
CC spearmint limonene synthase
XX

SQ Sequence 599 AA;

Alignment Scores:

Pred. No.: 2,75e-261 Length: 599
Score: 2825.00 Matches: 530
Percent Similarity: 99.08% Conservative: 7
Best Local Similarity: 97.79% Mismatches: 5
Query Match: 97.15% Indels: 0
DB: 3 Gaps: 0

US-09-938-956-6 (1-1632) x AAB29400 (1-599)

QY 4 AGAGATCCGGAAATACCAACCTTCTGTTGGATGTCACATTCATCAATCGGTTCTC 63
Db 58 ArgArgSerGlyAsnTyrAsnProSerArgTrpAspValGluPheIleGlnSerLeuHis 77
QY 64 AGTGACTATAAGGAGGACAAACACGCTGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
Db 78 SerAspTyrGluGluAspLysHisAlaIleArgAlaSerGluLeuValThrLeuValLys 97
QY 124 ATGGAACTGGAGAAAGAAACCGATCAAAATTCGACAACTTGATTCGATCGATTCGAG 183
Db 98 MetGluLeuGluLysGluThrAspHisIleArgGlnLeuGluLeuIleAspLeuGln 117
QY 184 AGGATGGGCTGCTCGCATATTCGAAATGATTCGAAAGATCTTCTCTCATATAT 243
Db 118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIleTyr 137
QY 244 CTCGACCATCACTATTACAAAGAACCTTTTCCAAAGAAAGGAGATCTCTACTCCACA 303
Db 138 LeuAspHisIleTyrTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157
QY 304 TCTCTTGCAATTTAGGCTCTCAGAGACATGTTTCAAGTCGCGCACAGAGTATTCGAT 363
Db 158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177
QY 364 AGTTTCAAGACGAGGAGGCTGAGTTCAAGAAAGCCTTAGCGACGACACACGAGGATG 423
Db 178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspThrLysGlyLeu 197
QY 424 TTGCAACTGTATGAACTCTCTTCTGTTGACGAGGCGAAACCAAGCTCGAGTCAGCG 483
Db 198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217
QY 484 AGGGAATTCGCCACCAAAATTTTGGAGGAAAGTGAACGAGGCTGTTGATCGCGAC 543
Db 218 ArgGluPheAlaThrLysPheLeuGluGluArgValAsnGluGlyGlyValAspGlyAsp 237
QY 544 CTTTAAACAGAAATCCATATCTTTGGAATCATCCCTTTCATGAGGATTAAGGCGCA 603
Db 238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgValLysArgPro 257
QY 604 AATGCACTCTGTGGATCGAATGTTATAGGAGGCGCGACATCAATCCAGTCTAGTTG 663
Db 258 AsnAlaProAlaTrpIleGluTrpTyrArgLysArgProAspMetAsnProValValLeu 277
QY 664 GAGCTTGCCATCTCGACTTAATATTTGTCAGACAAATTTCAAGACAAATTTCAAGAGAGTCAAGAA 723
Db 278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluLeuLysGlu 297
QY 724 TCCTTCAAGTGTGGAGAAATATCTGGTTTGTGTGAGAGCTCCCTTCGACGAGTAGA 783
Db 298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317
QY 784 CTGGTGGAAATGCTACTTTTGGAAATATCTGGGATCATCGAGCCACGTCAGCATCGCAAGTGCA 843

Db 318 LeuValGluCysTyrPheTrpAsnThrGlyIleLeuGluProArgGlnHisAlaSerAla 337
QY 844 AGGATATATGATGGGCAAGATCAACGCTCTGATTTACGGTATGATGATATTTATGATGTC 903
Db 338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357
QY 904 TATGGCACCTTAGAAGAACTCGAAACAATTCACCTGACCTCATTCGAAGATGGGATATAAC 963
Db 358 TyrGlyThrLeuGluGluLeuGluGlnPheThrGluLeuIleArgArgTrpAspIleAsp 377
QY 964 TCAATCGACCAACTTCCCGATATACATGCAACTGTGCTTTCTTCTGCACTCAACACTTCGTC 1023
Db 378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397
QY 1024 GATGATACATCGTACGATGTTATGAGAGGAGAAAGGCGTCAACCTTATACCTTACCTGCGG 1083
Db 398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417
QY 1084 CAATCTGGGTTGATTTGGGGATAGTATATGCTAGAGCACGCTGGTCTTCTACGCGGG 1143
Db 418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437
QY 1144 CACAAACCAAGTTTGAAGAGTATTTGGAGAACTCATGGCAGTCGATAGTGGGCTGT 1203
Db 438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457
QY 1204 ATGTTAACGACATATTTCTCCGAGTAACAGATTCGTTTCCAAAGGAGACCGTCGACAGT 1263
Db 458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477
QY 1264 TTGTACAAATACCAACGATTTAGTTCGTTGGTTCATCTCTGCTGGGCTTGTGTATGAT 1323
Db 478 LeuTyrLysTyrHisAspLeuValArgTrpSerPheValLeuArgLeuAlaAspAsp 497
QY 1324 TTGGAAACCTCGGTGAGAGGTGACGAGGGGATGTCGCGAAATCACTTCACTGCTAC 1383
Db 498 LeuGlyThrSerValGluGluValSerArgLysAspValProLysSerLeuGlnCysTyr 517
QY 1384 ATGAGTGACTACCAATCATCGAGGGGAGCGCGGAGCACCGTGAATGGCTGATAGCG 1443
Db 518 MetSerAspTyrAspAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla 537
QY 1444 GAGGTGGAAGAAAGATGAATCGGAGAGGCTGCAAGGATTCCTCATTCGCGCAAGAT 1503
Db 538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557
QY 1504 TTTTATAGGATGTCAGTTCGATTTTAGGAAGGATGCGCAGTTGATGTACCAATATGGAGAT 1563
Db 558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577
QY 1564 GGGCAGCGCACACCAACCTTATATACATCAACAAATGACCAAGACCTTATTCGAGCCC 1623
Db 578 GlyHisGlyThrGlnHisProIleHisGlnGlnMetThrArgThrLeuPheGluPro 597
QY 1624 TTTGCA 1629
Db 598 PheAla 599
RESULT 7
AAV02332
ID AAV02332 standard; protein; 591 AA.
XX
AC AAV02332;
XX
08-JUL-1999 (first entry)
XX
Computer generated 1,8-cineole synthase variant.
XX
(+)-bornyl diphosphate synthase; sage; 1,8-cineole synthase;
KW
(+)-sabinene synthase; monoterpene production; plant flavour;
KW
plant aroma; plant defence.
XX

QY 1552 CATATGAGATGGCGACGACACACACCTTATTATACATCAACAAATGACGACGAC 1611
 Db 566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585
 QY 1612 TTATTCGAGCCCTTT 1626
 Db 586 LeuPheAspArgTyr 590

RESULT 8
 AAY02335
 ID AAY02335 standard; protein; 591 AA.
 XX
 AC AAY02335;
 XX
 DT 08-JUL-1999 (first entry)
 XX
 DE Computer generated 1,8-cineole synthase variant.
 XX
 KW (+)-bornyl diphosphate synthase; sage; 1,8-cineole synthase;
 KW (+)-sabinene synthase; monoterpene production; plant flavour;
 KW plant aroma; plant defence.
 XX
 OS Synthetic.
 OS Salvia officinalis.
 XX
 XX W09915624-A1.
 XX
 PD 01-APR-1999.
 XX
 PF 25-SEP-1998; 98WO-US020120.
 XX
 PR 25-SEP-1997; 97US-00937540.
 XX
 PA (UNIW) UNIV WASHINGTON STATE RES FOUND.
 XX
 PI Croteau RB, Wise ML, Katahira EJ, Savage TJ;
 XX
 DR WPI; 1999-254698/21.
 XX
 XX Nucleic acid encoding monoterpene synthesis enzymes from sage, used e.g.
 PT to modify flavor, aroma, and defense capability.
 PT
 XX Example 7; Page 244-246; 259pp; English.
 PS
 CC AAY02329-38 represent computer generated variants of the enzyme 1,8-
 CC cineole synthase from sage. The specification also describes sage (+)-
 CC sabinene synthase and (+)-bornyl diphosphate synthase. Expression vectors
 CC containing nucleic acids encoding these enzymes are used to alter levels
 CC of the specified enzymes in host cells, to increase production of
 CC monoterpenoids, e.g. to modify flavour or aroma of plants, to improve
 CC their defence capability or to alter other interactions mediated by
 CC monoterpenes or produce the enzymes for subsequent use
 XX
 SQ Sequence 591 AA;

Alignment Scores:
 Pred. No.: 3-22e-147 Length: 591
 Score: 1634.50 Matches: 305
 Percent Similarity: 75.96% Conservative: 109
 Best Local Similarity: 55.96% Mismatches: 116
 Query Match: 56.21% Indels: 15
 DB: 2 Gaps: 7

US-09-938-956-6 (1-1632) x AAY02335 (1-591)

QY 4 AGAGATCCGGAACACACCTTCCTGCTGGGATGTCACCTTCATCCATCGCTTC 63
 Db 57 ArgArgThrGlyGlyTyrGlnProThrLeuTrpAspPheSerThrGlnLeuPheAsp 76
 QY 64 AGTGACTATAGAGAGGACAAACACGCTGATPAGGGCTTCGAGCTGCTCATTGGTGAAG 123
 Db 77 SerGluTyrLysGluGluLysHisLeuMetArgAlaGlyMetIleAlaGlnValAsn 96

124 ATCGAACTCGAGAAAGAAACCGATCAAAATTCGACAACTTGAGTTGATCGATGACCTGCAG 183
 Db MetValLeuGlnGluValAspSerIleGlnArgLeuGluLeuIleAspLeuArg 116
 QY 184 AGGATGGGCTGCTCCGATCATTTCCAAATAGTTCAAAGAAATCTTGCTCTATATAT 243
 Db ArgLeuGlyIleSerCysHisPheAspArgGluIleValGluLeuAsnSerIleTyr 136
 QY 244 CTGACCATCATCTATTACAAAGAACCCCTTTTCCAAAGAAAGAGGATCTCTACTCCACA 303
 Db TyThrAsnAsnGluIle-----AspGluSerAspLeuTyrSerThr 150
 QY 304 TCTCTTGCATTTAGGCTCTCTCAGAGAACATGGTTTTCAGAGTCGACCAAGAGGTATTCGAT 363
 Db AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp 170
 QY 364 AGTTTCAAGAAACGAGGAGGT---GAGTTCAAGAAAGCGCTTAGCGACGACACACGAGGA 420
 Db CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190
 QY 421 TTGTTCAACTGTATCAAGCTTCCTTCTGTCGAGGAAAGGCGAACCAGCTCGAGTCA 480
 Db LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu 210
 QY 481 GCGAGGGAATTCGCCACCAAAATTTTGGAGGAAAGTGAACGAGGCTGTGTGTATGGC 540
 Db AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu-----ValAspLys 227
 QY 541 GAC-----CTTTTAAACAAGATCGCATATCTTTTGACATCCCTCTTCTATTCGAGATT 594
 Db AspIleAsnLeuLeuSerSerIleGluArgAlaLeuGluLeuProThrHisTrpArgVal 247
 QY 595 AAAAGCCCAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCGACATGAATCCA 654
 Db GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267
 QY 655 GTAGTGTGGAGCTTGCCATCTGCTGCTGATGATGATTAATATTGTTCAAGCAACAATTTCAGAGAG 714
 Db ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlu 287
 QY 715 CTCAAGAAATCCTCAGGCTGCGAGAAATACCTGGGTGTTGTCGAGAGCTGCCCTTCGCA 774
 Db LeuLysGluAlaSerArgTrpTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307
 QY 775 AGGATAGACTGTGGAAATCTACTTTTGAATACTGGGATCATCGGACCCACCTCAGCAT 834
 Db ArgAspArgIleValGluCysTyrTyrThrThrGlyValValGluArgGluHis 327
 QY 835 GCAAGTCGAAGGATATGATGCGCAAGTCAACGCTCTGATTAAGGTCGATGATATT 894
 Db GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrThrIleAspAspVal 347
 QY 895 TATGATGCTATGCGACCTTAGAAGAACTCGAACAATTCATGACCTCATTCGAGAGATGG 954
 Db PheAspIleTyrGlyThrLeuGluLeuGlnLeuPheThrThrAlaIleGlnArgTrp 367
 QY 955 GATATAAACTCAATCGACCACTTCCTCCGATTAATGCAACTGCTGCTTCTTCTGCACTCAAC 1014
 Db AspIleGluSerMetLysGlnLeuProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387
 QY 1015 AACTTCGTCGATGATACATCGTACGATGTTATGAAGGAAAGAGCGTCAACGTTTATACCC 1074
 Db AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407
 QY 1075 TACTTCGGGCAATCGTGGTGTGATTTGCGGATAGATATATGTTAGGAGGACGGTGTTC 1134
 Db TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTrpTyr 427
 QY 1135 TACGCGGGCAACCAACGATTTGGAAGATGATTTGAGAGCTCATGCGAGCTCATGATAGT 1194
 Db TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447

Db	211	AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu-----ValaAspLys	227
Qy	541	GAC-----CTTTTAAACAAGAATCGCATATCTTTGGACATCCCTCTTATTGGAGATT	594
Db	228	AspIleAsnLeuLeuSerIleGluArgAlaLeuGluLeuProThrHisrPArgVal	247
Qy	595	AAAAGGCCAAATGCACCTGTGTGGATCGAATGTATAGGAAGGCCGACATGAATCCA	654
Db	248	GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro	267
Qy	655	GTAGTGTGGAGCTTCCCATCTACTCGACTTAAATATTGTTCAACGACAAATTTCAAGAAGAG	714
Db	268	ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlnGlu	287
Qy	715	CTCAAGAAATCCCTTCAGTGGTGAGAGAAATACTGGGTGTTGTCGAGAGTGCCTTCGCA	774
Db	288	LeuLysGluAlaSerArgTTrpAsnSerThrGlyLeuValHisGluLeuProPheVal	307
Qy	775	AGGGATAGACTGGTGGATCGTACTTTTGGAACTACTCGGATCATCGAGCCACGTCAGCAT	834
Db	308	ArgAspArgIleValGluCysTyrTyrTTrpThrThrGlyValGluArgArgGluHis	327
Qy	835	GCAAGTCGAAGGATAAATGATGGGCAAGTCAACGCTCTGATTACGGTGTATCGATGATTT	894
Db	328	GlyTyrGlnArgIleMetLeuThrLysIleAsnAlaLeuValThrThrIleAspAspVal	347
Qy	895	TATGATCTCATGGCACTTAGAAGAACTCGAACTACTGACCTCATCTCGAAGATGG	954
Db	348	PheAspIleTyrGlyThrLeuGluGluLeuGlnLeuPheThrThrAlaIleGlnArgTTrp	367
Qy	955	GATATAAATCAATTCGACCAACTCCCGATTACATGCAACTGTGCTTCTTCGACACTCAAC	1014
Db	368	AspIleGluSerMetLysGlnLeuProProTyrMetGlnIleCysTyrLeuAlaLeuPhe	387
Qy	1015	AACTTCCTCGATGATACATCGTACGATGTATAGGAAGAAAGCGTCACGTTATACCC	1074
Db	388	AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro	407
Qy	1075	TACCTGGCGCAATCGTGGTGTGATTTCGCGGATAAGTATATGGTAGAGGACACGCTGCTC	1134
Db	408	TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTTrpTyr	427
Qy	1135	TACGGCGGGCAACAAACCAAGTTTCGAAGATTTTGGAGAACTCATCGCAGTCGATAGT	1194
Db	428	TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly	447
Qy	1195	GGGCCCCGTATGTTAAAGCACAATTTCTCCGACTACAGATTCGTTCCACAAAGGAGACC	1254
Db	448	GlyIleProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluGluAsp	467
Qy	1255	GTCACACAGTTTGTAACAATACACGATTTAGTTCGTTGGTCATCCTTCGTTCTCGCGCTT	1314
Db	468	AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu	487
Qy	1315	GCTGATGATTTGGAAACCTCGGTGGAAGAGTGTAGCAGAGGGGATGTCCGGAATCACTT	1374
Db	488	AlaAspAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProLysSerVal	507
Qy	1375	CAGTGTCTACATGATGACTCAATATGTCATCGAGAGCGGAGCGCGGAAGCACGTGAATGG	1434
Db	508	GlnCysTyrMetAsnGluLysAsnAlaSerGluGluGluAlaArgGluHisValArgSer	527
Qy	1435	CTGATAGCGGAGGTGTGGAAGAGATGATCGGAGAGGGGTGTGCGAGGATTTCCCATTC	1494
Db	528	LeuIleAspGlnThrTrpLysMetMetAsnLysGluMetMetThr-----SerSerPhe	545
Qy	1495	GGCAAGATTTTATAGGATGTGCGATTCATTATTAGGAAGATGCGGCGAGTTGATGTAC---	1551
Db	546	SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln	565
Qy	1552	CATATGAGATGGCGACGGCACACAAACCCCTATTATACATCAACAAATGACCGAGAAC	1611

Db	566	HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu	588
Qy	1612	TTATTTCGAGCCCTTT	1626
		::: :::	
Db	586	LeuPheAspArgTyr	590
RESULT 11			
AAAY90839			
ID	AAAY90839	standard; protein; 591 AA.	
XX	AC		
XX	AAAY90839;		
DT	25-AUG-2000	(first entry)	
DE	Sage 1,8	cinsole synthase protein sequence SEQ ID NO:24.	
XX	Synthase;	protein co-ordinate data; active site; modification; terpenoid;	
XX	3-dimensional	coordinate; alpha carbon atom; plant; terpene synthase;	
KW	isoprenoid;	breeding programme; fragrance; flavour; pheromone;	
KW	defensive agent;	pigment; antitumour; steroid hormone;	
KW	signal transduction	pathway; bile acid; affinity purification;	
KW	photoreceptor;	enzymatic synthesis; nutrient supplement;	
KW	immunological	reagent.	
XX	Salvia officinalis.		
OS	WO200017327-A2.		
PN	30-MAR-2000.		
XX	17-SEP-1999;	99WO-US021419.	
XX	18-SEP-1998;	98US-0100993P.	
PR	22-APR-1999;	99US-0130628P.	
PR	23-AUG-1999;	99US-0150262P.	
XX	(KENT)	UNIV KENTUCKY RES DEPT.	
PA	(SALK)	SALK INST BIOLOGICAL STUDIES.	
XX	Chappell J,	Manna KR, Noel JP, Starks CM;	
XX	WPI;	2000-292839/25.	
DR	N-PSDB;	AAA38924.	
XX	Novel	terpene synthase enzymes, useful for producing terpene	
PT	hydrocarbons, e.g.	fragrances or antitumor agents, are derived from known	
PT	enzymes by	specific amino acid alterations.	
XX	Claim 114;	Page 376-377; 450pp; English.	
XX	The present	invention describes an isolated terpene synthase (I)	
CC	comprising a	region with at least 20% identity to region 265-535 of a 548	
CC	amino acid (aa)	sequence (Ia), given in AAAY90831. (I) contains nine alpha	
CC	-carbon atoms (alphaC)	that have interatomic distances, between each	
CC	other, within	tabulated ranges, have a centre point (within a sphere of	
CC	radius 2.3 Angstrom)	within tabulated ranges, and have an ordered	
CC	arrangement of R	groups (defining aa side chains), excluding specific	
CC	tabulated	arrangements (tables given in the specification). (I), and	
CC	related	enzymes, are used to produce a wide range of terpenoids (e.g.	
CC	cyclic, acyclic,	optionally hydroxylated), useful e.g. as fragrances,	
CC	flavours, pheromones,	defensive agents, pigments, antitumour agents,	
CC	components of	signal transduction pathways, precursors of steroid	
CC	hormones and	bile acids, as photoreceptors and as co-factor side chains.	
CC	Some	synthases with little or no catalytic activity (and nucleic acids	
CC	encoding them)	are used as controls in the analysis of products formed by	
CC	enzymatic	synthesis; as nutrient supplements; for affinity purification	
CC	of isoprenoids;	or to develop immunological reagents or nucleic acids for	
CC	monitoring	expression of terpene synthase or inheritance of the gene in	
CC	plant	breeding programs. The new synthases may produce novel terpene	
CC	products. AAA38910	to AAA38938 and AAAY90831 to AAAY90859 represent	
XX	sequences	used in the exemplification of the present invention	
XX	Sequence	591 AA;	

Alignment Scores:

Length:	4,028-147	591
Matches:	1633.50	305
Conservative:	75.96%	109
Mismatches:	55.96%	116
Indels:	56.17%	15
Gaps:	3	7
DB:		
Query Match:		
Best Local Similarity:		
Percent Similarity:		
Score:		
Pred. No.:		
Fragment: 000200		

US-09-938-956-6 (1-1632) x AAY90839 (1-591)

QY	4	AGAGGATCGGGAACCTACAAACCCCTTCTCGTTGGGATGTCAACTTCATCCGCTCTCTC	63
Db	57	A-gA-gThrGlyGlyTy-rGlnProThrLeuTrpAspPheSerThrIleGlnLeuPheAsp	76
QY	64	AGTGACTATAAGGAGGACAAACACGTGATTAGGCTTCTGAGCTGGTGCACTTTGGTGAAG	123
Db	77	SerGlutryLysGlnGluLysHisLeuMetArgAlaIleGlyMetIleAlaGlnValAsn	96
QY	124	ATGGAACCTCGAGAAAGAAACGGATCAAAATTCGACAACCTTGAGTTGATCGATGCTTCGAC	183
Db	97	MetLeuLeuGlnGluValAspSerIleGlnArgLeuGluLeuIleAspAspLeuArg	116
QY	184	AGGATGGGCTCTCGCATCTTCCAAATGAGTTCAAAGAAATCTGTCTCTATATAT	243
Db	117	ArgLeuGlyIleSerCysHisPheAspArgGluIleValGluLeuLeuAsnSerLysTrp	136
QY	244	CTCGACCATCACTATTACAGAACCCCTTTCCAAAAGAAAGAGGATCTCTACTCCACA	303
Db	137	-----TyrThrAsnAsnGluIle-----AspGluSerAspLeuTySerThr	150
QY	304	TCTTTGCAATTGAGCTCTCAGAAACATCGTTTTCAGATCGGCACAGAGGTATTTCGAT	363
Db	151	AlaLeuArgPheLysLeuLeuArgGlnTrpAspPheSerValSerGlnGluValPheAsp	170
QY	364	AGTTTCAAGAACGAGGAGGT---CAGTTTCAAAGAACCTTAGCGACGACACACAGAGA	420
Db	171	CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly	190
QY	421	TTGTTGCAACTGATGAAGCTTCCCTTCTGTGTGCGAAGCGCAACCCGCTCGAGTCA	480
Db	191	LeuLeuGlnLeuTy-GluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu	210
QY	481	GGGAGGAATTCGCCACCCAAATTTTGGAGAAAGAAAGTGAACGAGGCTGTGTTGATGCG	540
Db	211	AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu-----ValAspLys	227
QY	541	GAC-----CTTTTAACAGGAATCCCATATCTTTGGACATCCCTCTTCATTGGAGGAT	594
Db	228	AspIleAsnLeuLeuSerSerIleGluArgAlaLeuLeuLeuProThrHisLstTrpGVal	247
QY	595	AAAAGCCAAATGCACCTGTGTGATCGAATGGTATAGAGAGAGCCCGACATGAATCCA	654
Db	248	GlnMetProAsnAlaArgSerPheIleAspAlaTyrlsArgGProAspMetAsnPro	267
QY	655	GTAGCTTGGAGCTTGCCATCTCCACITTAATATTTCAAGCACATTTTCAAGAGAGAG	714
Db	268	ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlnGlu	287
QY	715	CTCAAAGAATCCCTTCAGGTGGTGGAGAATACTCGGTTTGTGTGAGAAGCTGCCCTTCGA	774
Db	288	LeuLysGluAlaSerArgTrpIleAsnSerThrGlyLeuValHisGluLeuProPheVal	307
QY	775	AGGATAGACTGGTGAATGCTACTTTTGGAAATCTCGGATCATCGAGCCACGCTCAGCAT	834
Db	308	ArgAspArgIleValGluCysTrpTrpThrThrGlyValValGluArgGlnGluHis	327
QY	835	GCAAGTGCAAGGATAATGATGGCAAGAGTCAACGCTCTGATTACGTTGATCGATGATATT	894
Db	328	GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrIleAspAspVal	347
QY	895	TATGATCTTATGGCACCTTAGAAGAACTCGAAACAAATTCATGACCTCATCTCGAAGATGG	954

348	PhaAspIleTyrGlyThrLeuGluGluLeuGlnPheThrThrAlaIleGlnArgTyr	367
955	GATATAAATCAATGACCAACATCTCCGATATACATGCAATCGTCTTCTTGCACTCAAC	1014
368	AspIleGluSerMetIysGlnLeuProPheTyrMetGlnIleCysTyrLeuAlaLeuPhe	387
1015	AACCTCGTCATCATACATCGTACGATGTATTGAAGAGAGAGCGCTCAACGTTATACCC	1074
388	AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspIysGlyPheAsnSerThrPro	407
1075	TACCTCGGCATCTCGGTGTGATTTGGCGGATTAAGTATATGGTAGAGCAGCGTGCTC	1134
408	TyrLeuArgIysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaIysTrpTyr	427
1135	TACGGCGGGCACAAAACCAAGTTTGGAGAGTAGTATTGGAGAACTCATCGCAGTCGATAAGT	1194
428	TyrMetGlyHisIysProSerLeuGluGluTyrMetLeuAsnSerTrpIleSerIleGly	447
1195	GGGCCCTGTATGTTAAGCACAATATCTCTCCAGTAGTAACAGATTCTGTCACAAAGGAGACC	1254
448	GlyIleProIleLeuSerHisLeuPheArgLeuThrAspSerIleGluGluGluAsp	467
1255	GTGCACAGATTGTGACAAATACCAACGATTTAGTTTCGTTGGTTCATCCTTCGTTCTCGCGCTT	1314
468	AlaGluSerMetHisIysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu	487
1315	GCTGATGATTTGGGAACCTCGTGGAGAGGTGAGCAGAGGGGATGTCCGAATCACATT	1374
488	AlaAspAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProIysSerVal	507
1375	CAGTGCTACATGAGTACTACATCAATGTCATCGAGCGGAGCGCGGAGACGCGTGAATGG	1434
508	GlnCysTyrMetAsnGluIysAsnAlaSerGluGluGluAlaArgGluHisValArgSer	527
1435	CTGATACGAGAGGTGTGGAGAGAGATGATCGGAGAGGGTGTGCAAGGATTCTCCATTC	1494
528	LeuIleAspGlnThrTrpIysMetMetAsnLysGluMetMetThr-----SerSerPhe	545
1495	GGCAAGATTTTATAGGATGTGCAAGTTGATTTAGGAAGGATGCGCAGGTTGATGTAC---	1551
546	SerIysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln	565
1552	CATATGAGATGGGCACGCGCACACACACCCCTATTATACATCAACAAATGACCAGAACC	1611
566	HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu	585
1612	TTATTTCGAGCGCCTTT 1626	
586	LeuPheAspArgTyr 590	
RESULT 12		
AAAY02336		
ID	AAAY02336 standard; protein; 591 AA.	
XX		
AC	AAAY02336;	
XX		
DT	08-JUL-1999 (first entry)	
XX		
DE	Computer generated 1,8-cineole synthase variant.	
XX		
KW	(+)-bornyl diphosphate synthase; sage; 1,8-cineole synthase;	
KW	(+)-sabinene synthase; monoterpenoid production; plant flavour;	
KW	plant aroma; plant defence.	
XX		
OS	Synthetic.	
OS	Salvia officinalis.	
XX		
FN	W09915624-A1.	
XX		
PD	01-APR-1999.	
XX		
PF	25-SEP-1998; 98WO-US020120.	
XX		

RESULT 12

RESOLUTION 12
AAY02336
ID AAY02336 standard; protein; 591 AA.

AA
AC
AA02336:XX
DT 08-JUL-1999 (first entry)

Computer generated 1,8-cineole synthase variant.

XX (+)-bornyl diphosphate synthase; sage; 1,8-cineole synthase;
KW (+)-sabinene synthase; monoterpenoid production; plant flavour;
KW plant aroma; plant defence.

OS Synthetic.

OS Salvia officinalis.

PN WO9915624-A1.

01-APR-1999.

25-SEP-1998; 98WO-US020120.

PR 25-SEP-1997; 97US-00937540.
XX (UNIW) UNIV WASHINGTON STATE RES FOUND.
XX Croteau RB, Wise ML, Katahira EJ, Savage TJ;
XX WPI; 1999-254698/21.
XX Nucleic acid encoding monoterpene synthase enzymes from sage, used e.g.
XX to modify flavor, aroma, and defense capability.
XX Example 7; Page 246-248; 259pp; English.
XX AAY02329-38 represent computer generated variants of the enzyme 1,8-
XX cineole synthase from sage. The specification also describes sage (+)-
XX sabinene synthase and (+)-bornyl diphosphate synthase. Expression vectors
XX containing nucleic acids encoding these enzymes are used to alter levels
XX of the specified enzymes in host cells, to increase production of
XX monoterpenoids, e.g. to modify flavour or aroma of plants, to improve
XX their defence capability or to alter other interactions mediated by
XX monoterpenes or produce the enzymes for subsequent use
XX
XX Sequence 591 AA;
XX
XX Alignment Scores:
XX Pred. No.: 5,01e-147 Length: 591
XX Score: 1632.50 Matches: 305
XX Percent Similarity: 75.96% Conservative: 109
XX Best Local Similarity: 55.96% Mismatches: 116
XX Query Match: 56.14% Indels: 15
XX DB: 2 Gaps: 7
XX
XX US-09-938-956-6 (1-1632) x AAY02336 (1-591)
XX 4 AGAGATCCGGAACACTACACCCCTTCCTGCTGGATGTCACATTCATCCATCGCTTCTC 63
XX 57 ArgarghrrGlyGlyTyrGlnProThrLeuTipAspPheSerThrIleGlnLeuPheAsp 76
XX 64 AGTGAATTAAGAGAGCAACACGCTGATTAGGCTTCTGAGCTGCTACCTTTGGTGAAG 123
XX 77 SerGluTyrLysGluGluLysHisLeuMetArgAlaAlaGlyMetIleAlaGlnValasn 96
XX 124 ATGGAATCGGAGAAGAAACCGATCAAAATTCGACAACTTCAGTTCGATGATGATCGAG 183
XX 97 MetLeuLeuGlnGluLeuLeuAspSerIleGlnArgLeuGluLeuLeuLeuAspLeuArg 116
XX 184 AGATGGGCTGTCGCGATCATTTCCAAATAGATTCAAGAAATCTTGCTCTATATAT 243
XX 117 ArgLeuGlyLysSerCysHisPheAspArgGluIleValGluIleLeuAsnSerLysTyr 136
XX 244 CTCGACCATCACTATTACAGAACCTTTCCAAAAGAGAAAGGATCTCTACTCCACA 303
XX 137 -----TyrThrAsnAsnGluIle-----AspGluSerAspLeuTyrSerThr 150
XX 304 TCTCTTCGATTTAGGCTCCCTCAGAGAACATGGTTTTCAGTCGCACAGAGGATTCGAT 363
XX 151 AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp 170
XX 364 AGTTTCAAGACGAGGAGGT---GAGTTCAAGAAAGCCTTAGCAGCAGCACAGGA 420
XX 171 CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190
XX 421 TTGTTGCAACTGTATGAAGCTTCTCTGTTGACGAGGCGAAACACCGCTCGAGTCA 480
XX 191 LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu 210
XX 481 GCGAGGAATTCGCCACAAATTTTTCGAGAAAGAAAGTGAACGAGGGTGTGTTGATGGC 540
XX 211 AlaArgAspPheAlaThrLysPheHisLysArgValLeu-----ValAspLys 227
XX 541 GAC-----CTTTTAAAGAAATCGCATATTCTTTGGACATCCCTCTTCATTGAGGATT 594
XX 228 AspIleAsnLeuLeuSerIleGluArgAlaLeuGluLeuProThrHisLysTrpArgVal 247

QY 595 AAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGCCCGACATGAATCCA 654
DB 248 GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267
QY 655 GPAGTGTGGAGCTTGCATATCGACTTAATATTTGTTCAAGCACAATTTCAAGAGAG 714
DB 268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGln 287
QY 715 CTCRAAGATCCTTCAGCTGGTGGAGAAATACTGGTGTGGTGGAGAGCTGCCCTTCGCA 774
DB 288 LeuLysGluAlaSerArgTrpAsnSerThrGlyLeuValHisGluLeuProIleVal 307
QY 775 AGGGATAGACTGGTGGAAATGCTTCTTTGGATCTCGGATCATCGAGCCACGTCAGCAT 834
DB 308 ArgAspArgIleValGluCysTyrTyrTrpThrGlyValValGluArgArgGluHis 327
QY 835 GCAAGTGCAGATTAATGATGGGCAAGTCAACCTCTGATTCGGTGCATGATGATATT 894
DB 328 GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrIleAspAspVal 347
QY 895 TATGATGTCTATGCGACCTTAGAAGAACTCGAACAATTCACCTGACCTCATTCGAGATGG 954
DB 348 PheAspIleTyrGlyThrLeuGluGluLeuGlnLeuPheThrAlaIleGlnArgTrp 367
QY 955 GATATAAATCAATCGACCAACTTCCCGATTCACGAACTGTCGAACTGTGCTTCTGCATCAAC 1014
DB 368 AspIleGluSerMetLysGlnLeuProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387
QY 1015 AACTTCGTCGATGATACATGTCATGCTTATGAGGAGAAAGGGTCAACGTTATATCCC 1074
DB 388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407
QY 1075 TACCTGCGCAATCGTGGTGTGATTTGGCGGATAGTATATGGTAGAGCGCGTGGTTC 1134
DB 408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTrpTyr 427
QY 1135 TACGCGCGGCAACAACCAAGTTTGGAGAGATTTTGGAGAACTCATGTCAGTCGATTAAGT 1194
DB 428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447
QY 1195 GGGCCCTCTATGTTAAACGCACATATTTCCGAGTAAACAGATTCGTTCCAAAGAGAGACC 1254
DB 448 GlyIleProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluAsp 467
QY 1255 GTCGACAGTTGTACAAATACACGATTTAGTTCGTTGCTCATCTCTGCTTCGCGCTT 1314
DB 468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487
QY 1315 GCTGATGATTTGGGAACCTCGGTGGAAGAGGTGACAGAGGGGATGTCGGAATTCACCTT 1374
DB 488 AlaAspAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProLysSerVal 507
QY 1375 CAGTCTCATGATGATGATCAATGTCATCGAGGCGGAGCGCGGAGCAGTGAATGG 1434
DB 508 GlnCysTyrMetAsnGluLysAsnAlaSerGluGluAlaArgGluHisValArgSer 527
QY 1435 CTGATAGCGGAGGTGTCGAAGAAGATGAATCGGAGAGGGTGTGCAAGGATTCCTCATTC 1494
DB 528 LeuIleAspGlnThrTrpLysMetMetAsnLysGluMetMetThr-----SerSerPhe 545
QY 1495 GCGAAAGATTTTATAGGATGTCGATTTGATTTAGGAAGATGCGCAGTGTGATGAC--- 1551
DB 546 SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln 565
QY 1552 CATATGAGATGCGCAGCGGACACACACCTATTATACATCAACAATGACCAAGAAC 1611
DB 566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585
QY 1612 TTATTCGAGCCCTTT 1626
DB 586 LeuPheAspArgTyr 590

RESULT 13
 AAY02329
 ID AAY02329 standard; protein; 591 AA.
 XX AC AAY02329;
 XX DT 08-JUL-1999 (first entry)
 XX DE Computer generated 1,8-cineole synthase variant.
 XX KW (+)-bornyl diphosphate synthase; sage; 1,8-cineole synthase;
 KW (+)-sabinene synthase; monoterpenoid production; plant flavour;
 KW plant aroma; plant defence.
 XX OS Synthetic.
 XX OS Salvia officinalis.
 XX PN WO9915624-A1.
 XX PD 01-APR-1999.
 XX PF 25-SEP-1998; 98WO-US020120.
 XX PR 25-SEP-1997; 97US-00937540.
 XX PA (UNIW) UNIV WASHINGTON STATE RES FOUND.
 XX PI Croteau RB, Wise ML, Katahira EU, Savage TJ;
 XX DR WPI; 1999-254698/21.
 XX PT Nucleic acid encoding monoterpene synthesis enzymes from sage, used e.g.
 PT to modify flavor, aroma, and defense capability.
 XX PS Example 7; Page 231-233; 259pp; English.
 XX CC AAY02329-38 represent computer generated variants of the enzyme 1,8-
 CC cineole synthase from sage. The specification also describes sage (+)-
 CC sabinene synthase and (+)-bornyl diphosphate synthase. Expression vectors
 CC containing nucleic acids encoding these enzymes are used to alter levels
 CC of the specified enzymes in host cells, to increase production of
 CC monoterpenoids, e.g. to modify flavour or aroma of plants, to improve
 CC their defence capability or to alter other interactions mediated by
 CC monoterpenes or produce the enzymes for subsequent use
 XX SQ Sequence 591 AA;

Alignment Scores:
 Pred. No.: 5,01e-147 Length: 591
 Score: 1632.50 Matches: 305
 Percent Similarity: 75.96% Conservative: 109
 Best Local Similarity: 55.96% Mismatches: 116
 Query Match: 56.14% Indels: 15
 DB: 2 Gaps: 7

US-09-938-956-6 (1-1632) x AAY02329 (1-591)

QY 4 AGAGGATCCGGAAGTACAAACCTTCTCGTGGGTGTCAACTTCATCCATCGCTTCTC 63
 Db 57 ArgAgtThrGlyGlyTyGlnProThrIlePheSerThrIleGlnLeuPheasp 76
 QY 64 AGTACTATAAGGAGGACAAACACGTTAGGGTCTCTGAGCTCGTCACTTTGGTGAAG 123
 Db 77 SerGluTyrLysGlnGluLysHisLeuMetAgaAlaGlyMetIleAlaGlnValAsn 96
 QY 124 ATCGAAGTGGAGAAAGACGATCAAAATCCGACAACTTGAGTGTGATCGATCGACTTGAG 183
 Db 97 MetLeuLeuGlnGlnGluValAspSerIleGlnArgLeuGlnLeuIleAspLeuArg 116
 QY 184 AGGATGGGCTGTCCGATCATTTCCAAAATGAGTTCAAGAATACTTGTCTCTATATAT 243
 Db 117 ArgLeuGlyIleSerCysHisPheAspArgGluIleValGluIleLeuAsnSerLysTyr 136

QY 244 CTCGACCATCACTATTACAAGAACCCCTTTTCCAAAGAAAGAAAGGATCTCTACTCCACA 303
 Db 137 -----TyrThrAsnAsnGluIle-----AspGluSerAspLeuTyrSerThr 150
 QY 304 TCTCTTGCAATTTAGGCTCTCAGAGAAACATGGTTTTCAAGTCGACACAGAGATATTCGAT 363
 Db 151 AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGlnValPheasp 170
 QY 364 AGTTTCAAGAACGAGGAGGGT---GAGTTTCAAGAAAGCCCTTTCAGCAGACACACGAGGA 420
 Db 171 CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190
 QY 421 TTGTTGCAACTGTATGAGCTTCCTTCTGTGTGCGAAGGCGGAACACACGCTCGAGTCA 480
 Db 191 LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu 210
 QY 481 GCGAGGGAATCCGCCACCAAAATTTTGGAGGAAAGTCAACGAGGGTGTGTTCATGGC 540
 Db 211 AlaArgAspPheAlaThrLysPheLeuHisLysAsgValLeu-----ValAspLys 227
 QY 541 GAC-----CTTTTAACAAGATCCCATATCTTTTGGACATCCCTTTCATTTGGAGGAT 594
 Db 228 AspIleAsnLeuLeuSerSerIleGluArgAlaLeuGluLeuProThrHisTyrArgVal 247
 QY 595 AAAAGGCCAAATGCCACCTGTGTGATCGAATCGTATAGGAAGAGCCCGACATGAATCCA 654
 Db 248 GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267
 QY 655 GTAGTGTGGAGCTTGCCTACTCGACTTAAATATATTTTCAAGCACAATTTCAAGAGAG 714
 Db 268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlnGlu 287
 QY 715 CTCAGAAGATCCTTCAGGTGGTGGAGAAATCTGGGTTGTGTGAGAAGTGGCCCTTCGCA 774
 Db 288 LeuLysGluAlaSerArgTyrPheAsnSerThrGlyLeuValHisGluLeuProPheVal 307
 QY 775 AGGATPAGACTGGTGAATGCTACTTTTGGAAATCTGGGATCTGGGATCATCGAGCCAGTCA 834
 Db 308 ArgAspArgIleValGluCysTyrTyrTrpThrGlyValValGluArgGluHis 327
 QY 835 GCAAGTGCAGATATATGATGGCAAGTCAAGCTCTGATTACGGTGTATTCATGATAT 894
 Db 328 GlyTyGluArgIleMetLeuThrLysIleAsnAlaLeuValThrThrIleAspVal 347
 QY 895 TATGATGCTCTATGGCACCTTAGAAGAACTCGAACAATTCAGTCACTCATTCGAAGATGG 954
 Db 348 PheAspIleTyrGlyThrLeuGluLeuGlnLeuPheThrThrAlaIleGlnArgTyr 367
 QY 955 GATATAAATCAATCCACCACTTCCGATATACATGCACTGTGCTTCTTTCACCTCAAC 1014
 Db 368 AspIleGluSerMetLysGlnLeuProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387
 QY 1015 AACTTCGTTCGATCATACATCATCGTACGATGTTTGAAGGAGAAAGCGCTCAACGTTATACC 1074
 Db 388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407
 QY 1075 TACTGCGCAATCGTGGTGTGATTTGGCGGATAGTATATATGATAGTACGACACCGTGTTC 1134
 Db 408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTyrTrp 427
 QY 1135 TAGCGGGGACAAACCAAGTTTGGAGAGATATTTGGAGAACTCATCGCAGTGCATAGTAA 1194
 Db 428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447
 QY 1195 GGGCCCTGTATTAACGACATATTTCTCCGAGTAAACAGATTCGTTCACAAGGAGACC 1254
 Db 448 GlyIleProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluGluAsp 467
 QY 1255 GTCGACAGTTTGTACAAATACACGATTTAGTTTCGTGTGGTCACTCTTCGTCTCGCGCTT 1314
 Db 468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487
 QY 1315 GCTGATGATTGGGAACCTCGGTGGAGAGGTGACGAGGGGATGTGCCGAATCACTT 1374

368 AspIleGluSerMetLysGlnLeuProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387
1015 AACTTCGTGCATGATACATCATCGATGTTATGAAGGAGAAAGCGGTCAAGTTATACCC 1074
388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407
1075 TACCTCGCGCAATCGTGGGTGATTTCGCGGATAAGTATATGATGAGGACGCGGTGC 1134
408 TyrLeuArgLysAlaTyrValAspLeuValGluSerTyrLeuIleGluAlaLysTyrTyr 427
1135 TACGCGGGGCACAAACCAAGTTTCGAAGATATTTCGAAGACTCATGCGCAGTTCGATAGT 1194
428 TyrMetGlyHisLysProSerLeuGluGlyTyrMetLysAsnSerTrpIleSerIleGly 447
1195 GGGCCCTGTATTAAACGCACATATCTCCGAGTAAACAGATTTCGTTACAAAGGAGACC 1254
448 GlyIleProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluAsp 467
1255 GTCGACAGTTGTACAATACACAGATTAGTTAGTTGGTGTGCATCCTTGCTTCGGCGTT 1314
468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487
1315 GCTGATGATTGGGAACCTCGTGGTGAAGAGGTGAGCAGAGGGGATGTCCGAATCACTT 1374
488 AlaAspAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProLysSerVal 507
1375 CAGTGTCAATCAGTCAGTCACTACATGTCATCCGAGCGGAGCGCGGAAGCACGTGAATGG 1434
508 GlnCysTyrMetAsnGluLysAsnAlaSerGluGluGluAlaArgGluHisValArgSer 527
1435 CTCGATACGGAGGTGTGGAAGAGTAAATGCGAGAGGGGTTCGAAGGATTCCTCATTC 1494
528 LeuIleAspGlnThrTrpLysMetMetAsnLysGluMetMetThr-----SerSerPhe 545
1495 GCGAAAGATTTTATAGGATGTGCAGTTGATTAGTATTAGGAAGGATGCGCAGTTGATGTAC--- 1551
546 SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln 565
1552 CATATGAGATGGGACGGCACACACCTTATTATACATCAACAATGACCAAGACC 1611
566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585
1612 TTATTCGAGCCCTTT 1626
586 LeuPheAspArgTyr 590

RESULT 15	QY	244	CTCGACCATCACTATTACAAAGAACCTTTTCCAAAGAGAAAGGATCTCTACTCCACA	300
AAAY02337	QY	137	-----TyrThrAsnAsnGluIle-----AspGluSerAspLeuTy-SerThr	150
AC	QY	304	TCCTCTGCAATTAGGCTCCTCAGAACATGGTTTTCAGTCGCACAAGAGGTATTCGAT	363
XX	QY	151	AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp	170
DT	QY	364	AGTTTCAAGAACGAGGAGGCT--GAGTTTCAAGAAAGCCCTTAGCGACGACACCAAGAGA	420
XX	QY	171	CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspThrArgGly	190
XX	QY	421	TTGTTGCCAACTGTATCAAGCTTCCTTCTGTTCGCGAAGCGCAACACCGCTCGAGTCA	480
XX	QY	191	LeuLeuGlnLeuTy-GluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu	210
OS	QY	481	GCGAGGGAATTCGCCACCAAAATTTTGGAGGAAAAGTCAAGAGGGGTGGTTCATGGC	540
XX	QY	211	AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu-----ValAspLys	227
PD	QY	541	GAC-----CTTTTAACAAGATCGCATATTCTTTGGACATCCCTCTTCATTGGAGGATT	594
XX	QY	228	AspIleAsnLeuLeuSerSerIleGluArgAlaLeuGluLeuProThrHisTrpArgVal	247
XX	QY	595	AAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGAGGCCCGACATGAATCCA	654
XX	QY	248	GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysAcArgProAspMetAsnPro	267
XX	QY		(UNIW) UNIV WASHINGTON STATE RES FOUND.	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame plus n2p model

Run on: February 25, 2004, 05:43:32 ; Search time 24 Seconds
(without alignments)
7021.132 Million cell updates/sec

Title: US-09-938-956-6
Perfect score: 2908
Sequence: 1 atgagacgattccgaaacta.....tattcgagcccttcacatga 1632

Scoring table: BLOSUM62

```

ygapop 10.0 , ygapext 0.5

```

devel	6.0	Devel	7.0
-------	-----	-------	-----

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 778828

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

```

-DB=/csm2.1/USPTO.spool
-PB=Issued Patents_AA
-OPWT=fastscan
-PATX=arai
-PA=24022004
-PA2=144702
-PA3=3089
-PA4=faster.fasta_1.179
-LOOPTXT=0
-UNITS=Bits
-START=1
-END=-1
-MATRIX=blosum62
-TRANS=human40.cdi
-LIST=45
-DLOCAL=200
-THR SCORE=pct
-THR MAX=100
-THR MIN=0
-ALIGN=15
-MODE=LOCAL
-OUTFMT=pct
-NORM=ext
-HEAPSIZE=500
-MINLEN=0
-MAXLEN=2000000000
-USER=US09938956
@CGN 1 1 27
@runat 24022004 144702 3089
-NCPU=6
-ICPU=3
-NO MMAP
-LARGEQUERY -NEG SCORES@0
-WAIT -DSFLOCK=100
-LONGLOG
-DBV TIMEOUT=120
-WARN
-TIMEOUT=30
-THREADS=1
-XGAPOP=1.0
-XGAPEXT=0.5
-FCGAPEXT=7
-XGAPOP=1.0
-YGAPEXT=0.5
-DELOP=6
-DEEXT=7

```

```

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB pep.*
5: /cgn2_6/ptodata/2/iaa/PTCTUS_COMB pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1 pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			ID	Description
	Score	Match	Length		
1	2876	98.9	599	2	US-08-846-526-11
2	2876	98.9	599	3	Sequence 11, Appli
3	2876	98.9	599	4	Sequence 2, Appli
4	2876	98.9	599	4	Sequence 22, Appl
5	2876	98.9	599	4	Sequence 22, Appl
6	2876	98.9	599	4	Sequence 22, Appl
7	2876	98.9	599	4	Sequence 22, Appl
8	1633.5	56.2	591	2	Sequence 4, Appli
9	1633.5	56.2	591	4	Sequence 24, Appl
10	1633.5	56.2	591	4	Sequence 24, Appl
11	1633.5	56.2	591	4	Sequence 24, Appl
12	1633.5	56.2	591	4	Sequence 24, Appl

1633.5	56.2	591	4	US-09-900-797-24	Sequence 24, Appl
1553.5	53.4	589	2	US-08-937-9540-6	Sequence 6, Appl
1553.5	53.4	590	4	US-09-398-398A-54	Sequence 54, Appl
1553.5	53.4	590	4	US-09-887-586A-54	Sequence 54, Appl
1553.5	53.4	590	4	US-09-887-586A-54	Sequence 54, Appl
1553.5	53.4	590	4	US-09-887-586A-54	Sequence 54, Appl
1553.5	53.4	590	4	US-09-903-012B-54	Sequence 54, Appl
1553.5	53.4	590	4	US-09-900-797-54	Sequence 54, Appl
1468	50.5	598	2	US-08-937-540-2	Sequence 2, Appl
1468	50.5	598	4	US-09-398-398A-36	Sequence 26, Appl
1468	50.5	598	4	US-09-887-586A-26	Sequence 26, Appl
1468	50.5	598	4	US-09-887-586A-26	Sequence 26, Appl
1468	50.5	598	4	US-09-895-753-26	Sequence 26, Appl
1468	50.5	598	4	US-09-903-012B-26	Sequence 26, Appl
1468	50.5	598	4	US-09-900-797-26	Sequence 26, Appl
1289.5	44.3	279	2	US-08-326-266-7	Sequence 7, Appl
836	28.7	556	4	US-09-398-398A-32	Sequence 32, Appl
836	28.7	556	4	US-09-887-586A-32	Sequence 32, Appl
836	28.7	556	4	US-09-887-586A-32	Sequence 32, Appl
836	28.7	556	4	US-09-895-753-32	Sequence 32, Appl
836	28.7	556	4	US-09-903-012B-32	Sequence 32, Appl
809.5	27.8	550	2	US-09-900-797-32	Sequence 8, Appl
809.5	27.8	550	2	US-08-443-633-8	Sequence 8, Appl
809.5	27.8	550	3	US-08-577-483-8	Sequence 8, Appl
809.5	27.8	550	4	US-09-435-380-8	Sequence 8, Appl
806.5	27.7	548	4	US-09-398-398A-8	Sequence 2, Appl
806.5	27.7	548	4	US-09-887-586A-2	Sequence 2, Appl
806.5	27.7	548	4	US-09-887-586A-2	Sequence 2, Appl
806.5	27.7	548	4	US-09-895-753-2	Sequence 2, Appl
806.5	27.7	548	4	US-09-903-012B-2	Sequence 2, Appl
806.5	27.7	548	4	US-09-900-797-2	Sequence 2, Appl
806.5	27.7	548	4	US-09-398-398A-8	Sequence 8, Appl
804.5	27.7	548	4	US-09-398-398A-12	Sequence 12, Appl
804.5	27.7	548	4	US-09-887-586A-8	Sequence 8, Appl
804.5	27.7	548	4	US-09-887-586A-8	Sequence 12, Appl
804.5	27.7	548	4	US-09-887-586A-8	Sequence 8, Appl
804.5	27.7	548	4	US-09-895-753-12	Sequence 12, Appl
804.5	27.7	548	4	US-09-895-753-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
 US-08-846-526-11
 Sequence 11, Application US/08846526
 Patent No. 5871988
 GENERAL INFORMATION:
 APPLICANT: Crouteau, Rodney B.
 APPLICANT: Colby, Sheila M.
 TITLE OF INVENTION: DNA Encoding Limonene Synthase From
 TITLE OF INVENTION: Mentha Spicata
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Christensen, O'Connor, Johnson & Kindness
 STREET: 2800 Pacific First Centre, 1420 Fifth Avenue
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98101-2347
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/846,526
 FILING DATE: 29-APR-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/582,802
 FILING DATE:
 APPLICATION NUMBER: US/08/145,941
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Shelton, Dennis K.
 REGISTRATION NUMBER: 26,997
 REFERENCE/DOCKET NUMBER: WSR-1-7219
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 224-0718
TELEFAX: (206) 224-0779
TELEX: 4938023
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 599 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-846-528-11

Alignment Scores:
Pred. No.: 2,528-284 Length: 599
Score: 2876.00 Matches: 542
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.90% Indels: 0
DB: 2 Gaps: 0

US-09-938-956-6 (1-1632) x US-08-846-526-11 (1-599)

```

QY 4 AGAGATCCGGAACTACCAACCTCTCGTGGATGTCAACTTCATCCAAATCGCTTCTC 63
DB 58 ArgSerGlyAsnTyrAsnProSerArgTrpAspValAsnPhelGlnSerLeuLeu 77
QY 64 AGTGACTATAAGGAGGACAAACACGTGATTAGGGCTTCAGCTGGTCTCACTTTGGTGAAG 123
DB 78 SerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuValLys 97
QY 124 ATCGAATCGAGAGAAAGCAAGTCAAAATCGCACTTGGTGTGATCGATCGATCGATCGAG 183
DB 98 MetGluLeuLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspAspLeuGln 117
QY 184 AGGATGGGGTGCCGATCATTTCCAAATAGATTCAAAAGAAATCTGTCTCTATATAT 243
DB 118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIleTyr 137
QY 244 CTCGACCATCATTTACAGAACCCCTTTTCCAAAGAGAAAGAGGATCTCTACTCCACA 303
DB 138 LeuAspHisIleTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157
QY 304 TCTCTTCATTTAGGCTCCAGAACATGGTTTTCAGTCGCAACAAGAGGTATTTCGAT 363
DB 158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177
QY 364 AGTTTCAAGAACGAGAGGGTGAGTTCAGAAAGCCCTTAGCGACACACCGAGATG 423
DB 178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197
QY 424 TTGCACTGTATGAAGCTCTCTTCTGTCAGCGAAGGCGAACACCGCTTCGAGTCAGCG 483
DB 198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217
QY 484 AGGGAATTCGCCACCAAAATTTTGGAGGAAAAAGTGAACGAGGTGGTGTGATGCGGAC 543
DB 218 ArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyValAspGlyAsp 237
QY 544 CTTTACAGAAATCGCATATTTCTTCGACATCCCTCTTCATTGGAGGATTTAAAGGCCA 603
DB 238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro 257
QY 604 AATGCACCTGTGTGATCGAATGATATAGGAAGGCGCCGACATGAATCGAGTAGTGTG 663
DB 258 AsnAlaProValTrpIleGluTyrArgLysArgProAspMetAsnProValValLeu 277
QY 664 GAGCTTCCGCTACTCGACTTAATATTGTTTCAAGCACAAATTTCAAGAGAGCTCAAGAA 723
DB 278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluLeuLysGlu 297
QY 724 TCCTTCAGGTGGTGGAGAAATCTCGGTTTGTGAGAAAGTGCCTTCGCAAGGGATAGA 783
DB 298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317

```

```

QY 784 CTGGTGAATGCTACTTTTGGAACTACTGGATCATCGAGCCACGTCAGATGCAAGTGCA 843
DB 318 LeuValGluCysIyrPheTrpAsnThrGlyIleGluProArgGlnHisAlaSerAla 337
QY 844 AGGATAATGATGGCAAGTCAACGCTCTGATTACGGTGATCGATGATATTATGATGTC 903
DB 338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357
QY 904 TATGGCACCTTAGAAGAACTCGCAACAAATTCACGTGACTCTATTCCGAAGATGGGATATAAAC 963
DB 358 TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgArgTrpAspIleAsn 377
QY 964 TCAATCGACCAACTTCCGATTAATGCAACTGTCTTTCTTGCATCAACAATCTCGTC 1023
DB 378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397
QY 1024 GATGATACATCGTACGATGTTATGAGGAGAAAGGGTCAACGTGATATACCTACCTCGCG 1083
DB 398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417
QY 1084 CAATCGTGGGTGATTGGCGGATAAGTATATATGATAGAGCACGCTGGTGTCTACGGCGG 1143
DB 418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437
QY 1144 CACAAACCAAGTTGGAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGT 1203
DB 438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457
QY 1204 ATGTAAACGACATATTTCCGAGTAAACAGATTCGTTCAAAAGGAGACCGTCGACAGT 1263
DB 458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477
QY 1264 TTGTACAATACCAAGATTTAGTTCGTTGCTCATCTCTGCTTCTCGCGCTTGTGTATGAT 1323
DB 478 LeuTyrLysTyrHisAspLeuValArgTrpSerPheValLeuArgLeuAlaAspAsp 497
QY 1324 TTGGAAACCTCGTGGAGAGGTGAGCAGAGGGGATGTCCGCAAAATCACTTCAGTGTCTAC 1383
DB 498 LeuGlyThrSerValGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517
QY 1384 ATGAGTGACTACAATGATCGAGCGGAGCGGAGCGGAGCAGCGTAATCGCTGATGATGCG 1443
DB 518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla 537
QY 1444 GAGGTGTGGAAGAGATGAATCGGAGAGGTGTGCAAGGATTCCTCATTCGGCAAAAGAT 1503
DB 538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557
QY 1504 TTTATAGGATGTCAGTGTGATTTAGGAAGGATGCGGAGTGTGATGATGATGATGATGAT 1563
DB 558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577
QY 1564 GGGCACGCGACACACACCTTATATACATCAACAAATGACCAAGACCTTATTCGAGCGCC 1623
DB 578 GlyHisGlyThrGlnHisProIleIleHisGlnMetThrArgThrLeuPheGluPro 597
QY 1624 TTGCA 1629
DB 598 PheAla 599

```

RESULT 2

US-09-172-339-2
Sequence 2, Application US/09172339
Patent No. 6291745
GENERAL INFORMATION:
APPLICANT: Euclidean Meyer, Terry
APPLICANT: Valpani, Nasser
TITLE OF INVENTION: Limonene and Other Downstream
Metabolites of Geranyl Pyrophosphate for Insect Control in
TITLE OF INVENTION: Plants
FILE REFERENCE: 5718-65
CURRENT APPLICATION NUMBER: US/09/172,339
CURRENT FILING DATE: 1998-10-14

EARLIER APPLICATION NUMBER: 08/449,061
EARLIER FILING DATE: 1995-05-24
EARLIER APPLICATION NUMBER: 08/153,544
EARLIER FILING DATE: 1993-11-16
EARLIER APPLICATION NUMBER: 08/042,199
EARLIER FILING DATE: 1993-04-02
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 599
TYPE: PRT
ORGANISM: Mentha spicata
US-09-172-339-2

Alignment Scores:

Pred. No.: 2,526-284 Length: 599
Score: 2876.00 Matches: 542
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.90% Indels: 0
DB: 3 Gaps: 0

US-09-938-956-6 (1-1632) x US-09-172-339-2 (1-599)

QY	4	AGACGACCGGAACTACGACCTCTCTGTTGGGATGTCAACTTCATCCATCGCTCTC	63
DB	58	ArgArgSerGlyAsnProSerArgTrpAspValAsnPhelIeGlnSerLeuLeu	77
QY	64	AGTGAATATAAGGAGGACAAACACGCTGATTAGGGCTCTCGAGCTGGTCACTTTGGTGAAG	123
DB	78	SerAspTyrLysGluAspLysHisValIleargAlaSerGluLeuValThrLeuValLys	97
QY	124	ATGGACCTCGGAAAGAAACGATCAATTCGACAACTTGATTCGATCGATCACTTCGAG	183
DB	98	MetGluLeuGluLysGluThrAspGlnIleargGlnLeuGluLeuLeuAspLeuGln	117
QY	184	AGGATGGGGCTCTCGATCATTTCCAAATCAGTTCAAAGAAATCTTCTCTCTATATAT	243
DB	118	ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIleTyr	137
QY	244	CTCGACCATCATATTACAAGAACCTTTTCCAAAGAAAGAAAGGATCTCTACTCCACA	303
DB	138	LeuAspHisIleTyrTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr	157
QY	304	TCTCTTGCAATTTAGGCTCTCTCAGACACATGGTTTCAAGTCGCAACAGAGATTTCCGAT	363
DB	158	SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp	177
QY	364	AGTTTCAAGAACGAGGAGGTGAGTTCAAAGAAAGCCTTTCAGCAGCACACAGAGGATG	423
DB	178	SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu	197
QY	424	TTGCACTGTATGAGCTTCTCTTCTGTTGCGGAGCGGAAACCACTCGAGTCAGCG	483
DB	198	LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrLeuGlnSerAla	217
QY	484	AGGGAATTCGCCACCAAAATTTTGGAGGAAAAGTGAACGAGGGTGTGTTGATGGCGAC	543
DB	218	ArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyValAspGlyAsp	237
QY	544	CTTTTAAAGAAATCGCATATTCTTTGGACATCCCTCTTCATTCGAGGATTAAGGCCA	603
DB	238	LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro	257
QY	604	AATGACCTGTGTGGATCAATGGTATAGGAAGCGCCGACATCAATCCAGTAGTGTG	663
DB	258	AsnAlaProValTrpIleGluTrpTyrArgLysArgProAspMetAsnProValValLeu	277
QY	664	GAGCTTGCCATCTCGATTAATATTTGTTCAAGCACAAATTTCAAGAGAGCTCAAGAA	723
DB	278	GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluGluLeuLysGlu	297
QY	724	TCCTTCAGGTGGTGAGAAATACTCGGTTTGTGTGAGAGCTGCCCTTCGCAAGGATAGA	783

DB	298	SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg	317
QY	784	CTGGTGGAAATGCTACTTTTGGAAATACCTGGATCATCGAGCCAGCTCAGCATGCAAGTGA	843
DB	318	LeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla	337
QY	844	AGGATAATGATGGGCAAACTCAACGCTCTGATTACCGTTCGATCGATGATATTATGATGTC	903
DB	338	ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal	357
QY	904	TATGGCACCTTGAAGAACTCGCAAACTCACTGACCTCATTCGAAGATGGGATATAAAC	963
DB	358	TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleargTrpAspIleAsn	377
QY	964	TCAATCGACCACTTCCCGATTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTC	1023
DB	378	SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnPheVal	397
QY	1024	GATGATACATCGTACGATGTTATGAGGAGAAAGGGCTCAACGTTTATACCTACCTGCGG	1083
DB	398	AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg	417
QY	1084	CAATCTGGGCTTGAATTTGGCGGATAAGTATATGTTAGAGGACAGCTGCTTCTACGGCGG	1143
DB	418	GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly	437
QY	1144	CACAAACCAAGTTTGAAGAGTATTTGGAGAACTCATGCGAGTCGATGATGAGTGGCCCTGT	1203
DB	438	HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys	457
QY	1204	ATGTTAAACGACATATTCTTCCAGTAAACAGATTCTGTTCAAAAGAGAGACCGTCGACAGT	1263
DB	458	MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer	477
QY	1264	TTGTAACAATFACCACTTTAGTTTGGTTCATCTCTGCTTCTGCGGCTTGTGATGAT	1323
DB	478	LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp	497
QY	1324	TTGCGAACCTCGGTGGAAGAGGTGACAGAGGGGATGTGCCGAAATCACTTTCAGTGTCTAC	1383
DB	498	LeuGlyThrSerValGluGluValSerArgLysValProLysSerLeuGlnCysTyr	517
QY	1384	ATGAGTGACTACAACTCATCGGAGGCGGAGCGGAGACACGCGTGAATGCTGATAGCG	1443
DB	518	MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla	537
QY	1444	GAGGTGGAAGAGAGTGAATGCGGAGAGGGTGTCCGAAGATTCCTCCATTCCGCAAGAT	1503
DB	538	GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp	557
QY	1504	TTTATAGGATGTGAGTTGATTTAGGAAGGATGGCGCAGTTGATGTACCAATATGGAGAT	1563
DB	558	PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp	577
QY	1564	GGGACGCGCACACACACCTTATACATCAACAATGACAGAACCTTATTCGAGCCC	1623
DB	578	GlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGluPro	597
QY	1624	TTTGCA 1629	
DB	598	PheAla 599	

RESULT 3

US-09-398-395A-22
Sequence 22, Application US/09398395A
Patent No. 6468772
GENERAL INFORMATION:
APPLICANT: Chappell, Joseph
APPLICANT: No. 64687721, Joseph P.
APPLICANT: Starks, Courtney M.
APPLICANT: Manna, Kathleen R.
TITLE OF INVENTION: SYNTHASES

```
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/398,395A
; CURRENT FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Mentha spicata
; US-09-398-395A-22

Alignment Scores:
Pred. No.: 2,52e-284 Length: 599
Score: 2876.00 Matches: 542
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.90% Indels: 0
DB: 4 Gaps: 0

US-09-938-956-6 (1-1632) x US-09-398-395A-22 (1-599)
QY 4 AGACGATCCGGAACACTACACCTTCTCGTTGGGATGTCAACTTCACTCCAAATCGCTTCTC 63
DB 58 ArgArgSerGlyAsnTyrAsnProSerArgTrpAspValAsnPheLeuGlnSerLeuLeu 77
QY 64 AGTGACTATAAGAGAGACAAACACGCTGATTAGGGCTTCTGAGCTGTCTCACTTTGGTGAAG 123
DB 78 SerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuValLys 97
QY 124 ATGGAACCTGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 183
DB 98 MetGluLeuLysGluThrAspGlnIleArgGlnLeuGluLeuLeuLeuLeuLeuLeuGln 117
QY 184 AGGATGGGGTGTCCGATCATTTCCAAATAGCTTCAAGAAATCTTGTCTCTATATAT 243
DB 118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluLeuSerSerIleTyr 137
QY 244 CTCGACCATCATATTACAGAACCTTTTCCAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 303
DB 138 LeuAspHisIleTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157
QY 304 TCTCTTGCAATTTAGCTTCTCAGACACATGTTTCAAGTCGACACAGAGGTATTTCGAT 363
DB 158 SerLeuAlaPheArgLeuLeuArgGluLysGlyPheGlnValAlaGlnGluValPheAsp 177
QY 364 AGTTTCAAGAACGAGGAGGTGAGTTCAAGAAAGCTTTAGCGACGACACACAGAGATTG 423
DB 178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspThrArgGlyLeu 197
QY 424 TTGCACTGTATGAGCTTCTTCTGTTGACGAGGCGGAACACGCTCCGAGTCAGCG 483
DB 198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrLeuGluSerAla 217
QY 484 AGGGAATTCGCCACCAATTTTGGAGAAAGAAAGTGAACGAGGCTGTTGATGTCGCGAC 543
DB 218 ArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyGlyValAspGlyAsp 237
QY 544 CTTTAAACAAGATCCGATATCTTTGGACATCCCTCTTCATTTGAGGATTAAGGCCA 603
DB 238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro 257
QY 604 AATGCACCTGTGATCGAATGTATAGGAAGGCGCCACATGATCGATGATGTTG 663
DB 258 AsnAlaProValThrIleGluThrTyrArgLysArgProAspMetAsnProValValLeu 277
QY 664 GAGCTTCCCATACCTGACTTAATATTTGTTCAAGACCAATTTCAAGAGAGCTCAAGAA 723
```

RESULT 4

```
US-09-887-586A-22
; Sequence 22, Application US/09887586A
; Patent No. 6495354
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: NO. 64953541, Joseph P.
```

```
DB 278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluLeuLysGlu 297
QY 724 TCCTTCAGGTGGTGGAGAAATATCGGTTGTTGTTGAGAGCTGCCCTTCGACAGGATAGA 783
DB 298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317
QY 784 CTGGTGAATGCTACTTTTGGAAATACTGGGATCATCGAGCCAGCTCAGCATGCAAGTGCA 843
DB 318 LeuValGluCysTyrPheTrpAsnThrGlyIleGluProArgGlnHisAlaSerAla 337
QY 844 AGGATAATGATGGGCAAGTCAACGCTCTGATTACGGTGTATCGATGATATTTATGATGTC 903
DB 338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357
QY 904 TATGGCACCTTAGAGAACTCGAACAATTCACCTGACTTCATTCGAAGATGGGATATAAAC 963
DB 358 TyrGlyThrLeuGluGluLeuGlnPheThrAspLeuIleArgTrpAspIleAsn 377
QY 964 TCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTCTTGCACTCAACAATTCGTC 1023
DB 378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397
QY 1024 GATGATACATCGTACGATGTTATGAAGGAGAAAGGCTGCAACGTATACCTTACCTCGG 1083
DB 398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417
QY 1084 CAATCGTGGTGTGTTGGCGGATAGTATATGATGAGCGACGCTGCTTCTACGGCGGG 1143
DB 418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437
QY 1144 CACAAACAAATTTGGAAAGATTTGGAGAACTCATCGGAGTCGATGATGGGCGCTGT 1203
DB 438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457
QY 1204 ATGTTAAACGACATATTTCCGAGTAAACAGATTCGTTCAAAAGAGAGACCGTCGACAGT 1263
DB 458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477
QY 1264 TTGTACAAATACCAACGATTTAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1323
DB 478 LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497
QY 1324 TTGGGAACCTCGGTTGAGAGGTGAGCAGAGGGGATGTCGCAATCACTTTCAGTGTCTAC 1383
DB 498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517
QY 1384 ATGAGTGACTACAAATGATCGAGCGGCGGAGCGGAGCAACGTGAATGCTGATAGCG 1443
DB 518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla 537
QY 1444 GAGGTGTGGAAGAGATCAATGCGAGAGGGTGTGGAAGGATTCCTCATTCCGCAAAAGAT 1503
DB 538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557
QY 1504 TTTTAGATGTGCAAGTTGATTTAGGAAGGATGCGGAGTGTGATGTATACCTAATATGAGAT 1563
DB 558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577
QY 1564 GGGACGCGCACACAACACCTTATATCATCAACAATGACCAAGCTTATTCGAGCCC 1623
DB 578 GlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGluPro 597
QY 1624 TTTGCA 1629
DB 598 PheAla 599
```

; APPLICANT: Starks, Courtney M.
 ; APPLICANT: Manna, Kathleen R.
 ; TITLE OF INVENTION: SYNTHASES
 ; FILE REFERENCE: 07678-025001
 ; CURRENT APPLICATION NUMBER: US/09/887,586A
 ; CURRENT FILING DATE: 2001-06-22
 ; PRIOR APPLICATION NUMBER: 09/398,395
 ; PRIOR FILING DATE: 1999-09-17
 ; PRIOR APPLICATION NUMBER: 60/130,628
 ; PRIOR FILING DATE: 1999-04-22
 ; PRIOR APPLICATION NUMBER: 60/150,262
 ; PRIOR FILING DATE: 1999-08-23
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 22
 ; TYPE: PRT
 ; LENGTH: 599
 ; ORGANISM: Mentha spicata
 ; ORGANISM: Mentha-586A-22
 ; US-09-887-586A-22

Alignment Scores:
 Pred. No.: 2,528-284 Length: 599
 Score: 2876.00 Matches: 542
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.90% Indels: 0
 DB: 4 Gaps: 0

US-09-938-956-6 (1-1632) x US-09-887-586A-22 (1-599)

QY 4 AGACGATCCGGAATACAAACCTTCCTCGTGGGATGTCAACTTCATCCAACTCGCTTC 63
 Db 58 ArgArgSerGlyAsnTyrAsnProSerArgTyrAspValAsnPheileGlnSerLeuLeu 77
 QY 64 AGTGACTATAAGGAGCAAAACACGTGATTAGGGCTTCGAGCTGGTCACTTTGGTGAAG 123
 Db 78 SerAspTyrLysGluAspLysHisValileArgAlaSerGluLeuValThruValLys 97
 QY 124 ATGGAACTCGGAGAAACCGATCAAAATTCACAACTTCAGTTGATCGATGACCTGCGAG 183
 Db 98 MetGluLeuGluLysGluThruAspGlnileArgGlnLeuGluLeuileAspLeuGln 117
 QY 184 AGATGGGCTGTCCGATCATTTCCAAATGATGTTCAAGAAATCTTGTCTCTATATAT 243
 Db 118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluileLeuSerSerileTyr 137
 QY 244 CTCGACCATCATTTACAGAACCTTTTCCAAAGAGAGAGGATCTCTACTCCACA 303
 Db 138 LeuAspHisHietTyrLysAsnProPheProLysGluLysGluLysGluLysGluLys 157
 QY 304 TCTCTTGCAATTTAGGCTCTCAGAGAACATGTTTCAAGTCGCAAGAGGTATTCGAT 363
 Db 158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177
 QY 364 AGTTTCAAGAACGAGAGGAGTTCAGAAAGAACCTTAGCGACGACACAGAGGATTG 423
 Db 178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197
 QY 424 TTGCAACTGTATGACCTTCTCTGTTGACGAGGCGAAACCAACGCTCGAGTCAGCG 483
 Db 198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThruGluGlyGluThruLeuGluSerAla 217
 QY 484 AGGAAATCCGACCAAAATTTTGGAGGAAAAGTGAACGAGGGTGGTGTGATGGCGAC 543
 Db 218 ArgGluPheAlaThrLysPheLeuGluLysValAsnGluGlyGlyValAspGlyAsp 237
 QY 544 CTTTAAACAGATCGCATATCTTGGACATCCCTCTTCATTGGAGGATTAAGAGCCA 603
 Db 238 LeuLeuThrArgileAlaTyrSerLeuAspilleProLeuHietTyrPargilleLysArgPro 257
 QY 604 AATGCACTGTGTGGATCGAATGGTATAGGAAGCGCCGACATCAATCCAGTAGTGTG 663
 Db 258 AsnAlaProValTyrileGluTyrArgLysArgProAspMetAsnProValValLeu 277

RESULT 5

US-09-895-752-22

; Sequence 22, Application US/09895752

; Patent No. 6559297

QY 664 GAGCTTGCCATCTCGACTTTAAATATTGTTCAAGCAAAATTTCAAGAAGAGCTCAAGAA 723
 Db 278 GluLeuAlaileLeuAspLeuAsnileValGlnAlaGlnPheGlnGluLeuLysGlu 297
 QY 724 TCCTTCAGGTGGTGGAAATACCTGGTTTGTGAGAGCTGCCCTTCGCAAGGGATAGA 763
 Db 298 SerPheArgTyrPargAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317
 QY 784 CTGGTGAATGCTACTTTTGGAACTACTCGGATCATCGAGCCACGTCAGCATGCAAGTGCA 843
 Db 318 LeuValGluCysTyrPheTrpAsnThrGlyileGluProArgGlnHisAlaSerAla 337
 QY 844 AGGATATGATGGCAAGTCAACGCTCTGATTCAGGTGATCCGATGATTCATGATGTC 903
 Db 338 ArgileMetMetGlyLysValAsnAlaLeuileThrValileAspilleTyrAspVal 357
 QY 904 TATGGCACTTAGAAGAACTCGAAACAAATTCACCTGACCTATTCCGAAGATGGGATATAAAC 963
 Db 358 TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuileArgArgTyrAspilleAsn 377
 QY 964 TCAATCGAACCACTTCCGATTCATGCAACTGCTGCTTCTTTCGACTCAACACTTCGTC 1023
 Db 378 SerileAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397
 QY 1024 GATGATACATCGTACGATGTTATGAAGGAGAAAGGCTCAACGTATACCTTACCTTCGCG 1083
 Db 398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValileProTyrLeuArg 417
 QY 1084 CAATCGTGGTTCATTGGCGGATAAGTATATCGTAGAGCAAGTGGTTCAGCGCGG 1143
 Db 418 GlnSerTyrValAspLeuAlaAspLysTyrMetValGluAlaArgTyrPheTyrGlyGly 437
 QY 1144 CACAAACCAAGTTTGGAGAGTATTGGAGAACTCATGGCAGTCGATAAGTGGGCCCTGT 1203
 Db 438 HisLysProSerLeuGluGluLysLeuLysAsnSerTyrGlnSerileSerGlyProCys 457
 QY 1204 ATGTTTAAACGACATATTTCTTCCGAGTAAACAGATTCGTTTCAAAAGAGAGACCGTCGACAGT 1263
 Db 458 MetLeuThrHisilePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477
 QY 1264 TTGTACAATACCAACGATTTAGTTCGTTGCTATCTCTGCTTCGCGCTTGTGATGAT 1323
 Db 478 LeuTyrLysTyrHisAspLeuValArgTyrPheValLeuArgLeuAlaAspAsp 497
 QY 1324 TTGGAACTTCGTTGAGAGAGTTCAGCAGAGGGGATGTCGGAATTCACATTCAGTGCTAC 1383
 Db 498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517
 QY 1384 ATGAGTCACTAATGATCGAGGCGGAGGCGCGGAGCAGCAGTGAATCGGTAGCG 1443
 Db 518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTyrLeuileAla 537
 QY 1444 GAGTGTGGAAGAGATGAATGCGGAGAGGGTGTCCGAAGATTCTCCATTCCGCAAGAT 1503
 Db 538 GluValTyrLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557
 QY 1504 TTTATAGGATGTGAGTGTGATTAGGAAGGATCGGCGAGTGTGATTCATATATGAGAT 1563
 Db 558 PheileGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577
 QY 1564 GGGCAGCGCACACCAACCTTATATACATCAACAATGACCAAGAACCTTATTTCGAGCCC 1623
 Db 578 GlyHisGlyThrGlnHisProileileHisGlnGlnMetThrArgThrLeuPheGluPro 597
 QY 1624 TTTGCA 1629
 Db 598 PheAla 599

```

; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 65592971, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/895,752
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Mentha spicata
US-09-895-752-22

Alignment Scores:
Pred. No.: 2,52e-284 Length: 599
Score: 2876.00 Matches: 542
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.90% Indels: 0
DB: Gaps: 4

US-09-938-956-6 (1-1632) x US-09-895-752-22 (1-599)
QY 4 AGACGATCCGGAACATACAAACCTTCTCGTTGGGATGTCACCTTCATCGAATCGGTTCTC 63
DB 58 ArgArgSerGlyAsnTyrAsnProSerArgTyrAspValAsnPheileGlnSerLeuLeu 77
QY 64 AGTGACTATAAGGAGGACAAACACGATGATTAGGGCTTCTGAGCTGCTCACTTTGGTGAAG 123
DB 78 SerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuValLys 97
QY 124 ATGGAACCTCGAGAAAGAAACGATCAAAATTCGACAACTTGATGATCGATCGATCGAG 183
DB 98 MetGluLeuGluLysGluThrAspGlnIleArgGlnLeuLeuLeuIleAspAspLeuGln 117
QY 184 AGGATGGGCTGCTCCGATCATTTCCAAATGAGTTCAAGAAATCTTGCTCTATATAT 243
DB 118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluLeuSerSerIleTyr 137
QY 244 CTCGACCATCATATTACAGAACCCCTTTTCCAAAGAAAGAAAGGATCTCTACTCCACA 303
DB 138 LeuAspHisHisTyrTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157
QY 304 TCTCTTGCAATTAGGCTCCTCAGAGAACATGTTTCAAGTCGACACAGAGGTATTCGAT 363
DB 158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177
QY 364 AGTTTCAAGAACAGAGAGGTGAGTTCAAAGAAAGCTTTAGCGACGACACAGAGGATTTG 423
DB 178 SerPheLysAsnGlnGluGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197
QY 424 TTGCAACTGATGAAAGTTCCTTTCTGTTGACGGAAAGGCGAAACACGCTCAGTCAGCGG 483
DB 198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217
QY 484 AGGGAATTCGCCACCAAAATTTTGGAGGAAAGTGAACGAGGGTGGTGTTCATGTCGCGAC 543
DB 218 ArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyValAspGlyAsp 237
QY 544 CTTTAAACAGAAATCGCATATCTTTGGACATCCCTCTTCATTCGGAGGATTAAGGCGCA 603

```

```

238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro 257
604 AATGCACCTGTGTGGATCGAATGGTATAGGAAGAGCCCGCATGATGAATCGATGTGTG 663
258 AsnAlaProValTrpIleGluTrpTyrArgLysArgProAspMetAsnProValValLeu 277
664 GAGCTTGGCCATCTCCGACTTAATATTTTCAAGCACAATTTCAAGAGAGCTCAAGAA 723
278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGlnLeuLysGlu 297
724 TCCTTCAGGTGGTGGAGAAATACCTGGTTTGTGTGAGAACTGCGCCCTCCAGAGGTAGA 783
298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317
784 CTGCTGGAATGCTACTTTTGGATATCTGGATCATCGAGCCACGTCAGATGCAAGTCGA 843
318 LeuValGluCysTyrPheTrpAsnThrGlyIleGlnGluProArgGlnHisAlaSerAla 337
844 AGGATAATGATGGGCAAACTCAACGCTCTGATTACGGTATCGATGATATTTATGATGTC 903
338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357
904 TATGGCACCTTAGAAGACTCGAAGCAATTCAGCTGACCTCATTCGAAGATGGGATATAAAC 963
358 TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgArgTrpAspIleAsn 377
964 TCAATCGACCACTTCCGATTCATACGCAACTGTGCTTTCTTGCACCTCAACAACTTCGTC 1023
378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397
1024 GATGATACATCGTACGATGTTATGAGGAGAAAGGGCTCAACGTTTATACCTTACCTCGCG 1083
398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417
1084 CAATCGTGGTTCGATTTGGCGGATAGTATATGCTAGAGACGCTGGTGTCTACGGCGGG 1143
418 GluSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437
1144 CACAAACCAAGTTTGGAGAGTATTTGGAGAACTATGCGAGTCAATGCGAGTCGATAGTGGCCCTGT 1203
438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457
1204 ATGTTAACGACATATTTCCGAGTAAACAGATTCGTTCAACAAGAGACCGTCGACAGT 1263
458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477
1264 TTGTACAAATACCAACGATTTAGTTGCTGCTCATCTTCTGCGGCTGCTGATGAT 1323
478 LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497
1324 TTGGGAACCTCGTGGAGAGGTGACGAGAGGGATGCGCGAAATCACTTCAGTGTCTAC 1383
498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517
1384 ATGAGTCACTACATGATCGAGCGGAGCGGAGCGGAGCACGTAATCGCTGATAGCG 1443
518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla 537
1444 GAGTGTGGAGAGATGAATCGGAGAGGTGCGAGAGGTGCGAGGATTCCTCAATTCGCAAGAT 1503
538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557
1504 TTTATAGATGTGCAAGTTGATTAGGAAGGATGCGCAGTTGATGTACCATTAATGAGAT 1563
558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577
1564 GGGCACGCGCACAAACCTTATATACATCAACAATGACAGCAACCTTATTCGAGCCC 1623
578 GlyHisGlyThrGlnHisProIleHisGlnGlnMetThrArgThrLeuPheGluPro 597
1624 TTTTCCA 1629
598 PheAla 599

```

RESULT 6

US-09-903-012B-22
; Sequence 22, Application US/09903012B
; Patent No. 6569556
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 6569556, Joseph P.
; APPLICANT: Starks, Courtney R.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012B
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Mentha spicata
US-09-903-012B-22

Alignment Scores:
Pred. No.: 2,52e-284 Length: 599
Score: 2876.00 Matches: 542
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.90% Indels: 0
DB: 4 Gaps: 0

US-09-938-956-6 (1-1632) x US-09-903-012B-22 (1-599)

QY	4	AGAGCATCGGAACTACAAACCTCTCGTTGGGATGTCACTTCATCCATCGCTCTC	63
DB	58	ArgArgSerGlyAsnTyAsnProSerArgTrpAspValAsnPhelieGlnSerLeuLeu	77
QY	64	AGTGAATAAGGAGGACAAACACAGTATTAGGCGCTTCGAGCTGGTCACTTTGGTGAAG	123
DB	78	SerAspTyLysGluAspLysHisValileArgAlaSerGluLeuValThrLeuValLys	97
QY	124	ATGGAACTGGAGAAAGAACCGATCAATTCGACAACTTGAGTTGATCGATTCGAG	183
DB	98	MetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspAspLeuGln	117
QY	184	AGGATGGGGCTGCTCGGATCAATTCGAAATCAGTTCAAGAAATCTTGCTCTATATAT	243
DB	118	ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIleTy	137
QY	244	CTCGACCATCACTATTACAGAACCTTTTCCAAAAGAAAGAAAGGATCTCTACTCCACA	303
DB	138	LeuAspHisHisTyTyLysAsnProPheProLysGluGluArgAspLeuTySerThr	157
QY	304	TCTCTTGCACTTAGCTCTCAGAGAACATGGTTTCAAGTCGACAGAGGATTCGAT	363
DB	158	SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp	177
QY	364	AGTTTCAAGACGAGGAGGGTGAAGTTCAGAAAGCCCTTAGCGACGACACAGAGGATTG	423
DB	178	SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu	197
QY	424	TTGCAACTGTATGAAGCTCTCTTCTGTGACGAGGCGGAAACCACTCGAGTCAGCG	483
DB	198	LeuGlnLeuTyGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla	217
QY	484	AGGGAAATTCGCCACCAAAATTTTGGAGGAAAGTGAACGAGGGTGGTGTGATGGCGAC	543
DB	218	ArgGluPheAlaThrLysPheLeuGluLysValAsnGluGlyGlyValAspGlyAsp	237

QY	544	CTTTTAAACAAGAAATCGCATATTTCTTTGGACATCCCTCTTCATTGGAGGATTAAAGGCCA	603
DB	238	LeuLeuThrArgIleAlaTySerLeuAspIleProLeuHisTrpArgIleLysArgPro	257
QY	604	ATGCACTCTGTGTGATCGAATGGTATAGGAGAGCCCGACATCAATCCAGTAGTTTG	663
DB	258	AsnAlaProValTrpIleGluTrpTyArgLysArgProAspMetAsnProValValLeu	277
QY	664	GAGCTTGCCTACTCGACTTAATATTTGTTCAAGCACAAATTTCAAGAGAGCTCAAGAA	723
DB	278	GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluGluLeuLysGlu	297
QY	724	TCCTTCAGGTGTGAGAAATACCTGGTTGTTGAGAGCTGCCCTTCCAGGATAGTA	783
DB	298	SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg	317
QY	784	CTGGTGAATGCTACTTTTGGAAATCTGGGATCATCGAGCCACGTCAGCATGCAAGTGCA	843
DB	318	LeuValGluCysTyPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla	337
QY	844	AGGATAATGATGGCAAACTCAACGCTCTGATTCAGGTGATCGATGATATTTATGATGC	903
DB	338	ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyAspVal	357
QY	904	TATGGCACTTAGAAGAACTCGAACAATTCACGTCACCTCATTGAGAGATGGGATATAAC	963
DB	358	TyGlyThrLeuGluGluLeuGlnPheThrAspLeuIleArgArgTrpAspIleAsn	377
QY	964	TCAATCGACCAACTTCCCGATTACATGCAACTGTGCTTTCTTCGCACTCAACACTTCGTC	1023
DB	378	SerIleAspGlnLeuProAspTyMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal	397
QY	1024	GATGATACATCGTACGATGTTATGAGGAGAAAGGGCTCAAGTTATACCTTACCTCGCG	1083
DB	398	AspAspThrSerTyAspValMetLysGluLysGlyValAsnValIleProTyLeuArg	417
QY	1084	CAATCGTGGTGTGATTTGGCGGATAAGTATATGAGAGCGACGGTGTCTACCGCGGG	1143
DB	418	GlnSerTrpValAspLeuAlaAspLysTyMetValGluAlaArgTrpPheTyGlyGly	437
QY	1144	CACAAACCAAGTTTGAAGAGTATTGGAGAACTCATGTCAGTCTGATAGTGGCCCTGT	1203
DB	438	HisLysProSerLeuGluGluTyLeuGluAsnSerTrpGlnSerIleSerGlyProCys	457
QY	1204	ATGTTAACCCACATATTTCTCCGAGTAACAGATTCTGTTCCACAAAGAGAGACCGTCAG	1263
DB	458	MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer	477
QY	1264	TTGTACAAATACCCAGATTTAGTTGCTTGGTTCATCTCTGCTCGGCTTGTCTGATGAT	1323
DB	478	LeuTyLysTyHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp	497
QY	1324	TTGGGAACCTCGGTGGAAGAGGTGACGAGAGGGATGTCGCGAAATCACTTCAGTGTCTAC	1383
DB	498	LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTy	517
QY	1384	ATGAGTCACTCAATTCATCGAGGCGGAGCGCGGAGACACGTCGTAATGGCTGATAGCG	1443
DB	518	MetSerAspTyAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla	537
QY	1444	GAGGTGGAGAGAGATGATTCGCGAGAGGTGTCGAGGATTCCTCCATTCCGCAAGAT	1503
DB	538	GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp	557
QY	1504	TTTATAGGATGTCAGTTGATTTAGGAAGGATGGCGGCTTGTATGATGATGAGAT	1563
DB	558	PhelIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyHisAsnGlyAsp	577
QY	1564	GGCGACGGCACACACACCTTATATACATCAACAAATGACCGAGACCTTATTCGAGCCC	1623
DB	578	GlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGluPro	597

QY 1624 TTGCA 1629
 Db 598 Pheala 599

RESULT 7

US-09-900-797-22
 ; Sequence 22, Application US/09900797
 ; Patent No. 6645762
 ; GENERAL INFORMATION:
 ; APPLICANT: Chappell, Joseph
 ; APPLICANT: No. 66457621, Joseph P.
 ; APPLICANT: Starks, Courtney M.
 ; APPLICANT: Manna, Kathleen R.
 ; TITLE OF INVENTION: SYNTHASES
 ; FILE REFERENCE: 07678-025001
 ; CURRENT APPLICATION NUMBER: US/09/900,797
 ; CURRENT FILING DATE: 2001-07-06
 ; PRIOR APPLICATION NUMBER: US/09/398,395
 ; PRIOR FILING DATE: 1999-09-17
 ; PRIOR APPLICATION NUMBER: 60/130,628
 ; PRIOR FILING DATE: 1999-04-22
 ; PRIOR APPLICATION NUMBER: 60/150,262
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 22
 ; LENGTH: 599
 ; TYPE: PRT
 ; ORGANISM: Mentha spicata
 US-09-900-797-22

Alignment Scores:

Pred. No.: 2,52e-284 Length: 599
 Score: 2876.00 Matches: 542
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.90% Indels: 0
 DB: 4 Gaps: 0

US-09-938-956-6 (1-1632) x US-09-900-797-22 (1-599)

QY 4 AGAGATCCGGAACACTACACCCCTCTCGTTGGATGTCAACTCATCAATCGCTTCTC 63
 Db 58 ArgArgSerGlyAsnTyAsnProSerArgTTPAspValAsnPheIleGlnSerLeuLeu 77
 QY 64 AGTCACTATAGGAGGCAACACGCTGATTAGGCTTCTGAGCTGTCTCACTTTGTGAAG 123
 Db 78 SerAspTyLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuValLys 97
 QY 124 ATGGAACCTGGAAGAAACGAGTCAAAATTCGACAACTGAGTTGATCGATCGAG 183
 Db 98 MetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspAspLeuGln 117
 QY 184 AGGATGGGCTGCTCGATCATTTCCAAATGAGTTCAAAGAAATCTTGCTCTATATAT 243
 Db 118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluLeuSerSerIleTy 137
 QY 244 CTCGACCATCATATTACAGAACCCCTTTCCAAAGAAAGAAAGGATCTTACTCCACA 303
 Db 138 LeuAspHisIstTyTyLysAsnProPheProLysGluLysGluLysLeuTySerThr 157
 QY 304 TCTCTTGATTTAGCTTCTCAGAGAACATGTTTCAAGTCGACCAAGAGGTATTTCGAT 363
 Db 158 SerLeuAlaPheArgSerLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177
 QY 364 AGTTTCAAGAACGAGGAGGTGAGTTCAAGAAAGCCCTTAGCGACGACACGAGATTG 423
 Db 178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197
 QY 424 TTGCAACTGTATGAAGTCTCTTCTGTCACGGAAGGCGAACACCGCTTCAGTTCAGCG 483
 Db 198 LeuGlnLeuTyGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217

QY 484 AGGAAATTCGCCACCAAAATTTTGGAGGAAAAAGTGAACAGGGTGTGTGTTGATGGCAG 543
 Db 218 ArgGluPheAlaThrLysPheLeuGluLysValAsnGluGlyValAspGlyAsp 237
 QY 544 CTTTAAACAAGATCGCATATTTTGGACATCCCTCTTTCATTGGAGGATTAAGAGGCA 603
 Db 238 LeuLeuThrArgIleAlaTySerLeuAspIleProLeuHisTrpArgIleLysArgPro 257
 QY 604 AATGCACCTGTGTGGATCGAATGTATAGGAAGAGCCCGACATGAATCCAGTAGTGTG 663
 Db 258 AsnAlaProValTrpIleGluTrpTyArgLysArgProAspMetAsnProValValLeu 277
 QY 664 GAGCTTGCATCTCGACCTTAATATTTCAAGCACAAATTTCAAGAAAGAGCTCAAGAA 723
 Db 278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluLeuLysGlu 297
 QY 724 TCCTTCAGGTGTGGAGAAATCTGGTTTGTGTAAGAGTCCCTTCGCAAGGATAGA 783
 Db 298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317
 QY 784 CTGCTGGAATCTACTTTTGGAAATCTGGATCATCGAGCCACGTCAGCATGCAAGTGA 843
 Db 318 LeuValGluCysTyPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla 337
 QY 844 AGGATATGATGGGCAAGCTCAACCTCTGTATTCGTTGATCGATGATATTTATGATGC 903
 Db 338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyAspVal 357
 QY 904 TATGGCACCTTAGAAGAACTCGAACCAATCTAGTCACTCATTCGAAAGATGGCATATAAC 963
 Db 358 TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgArgTrpAspIleAsn 377
 QY 964 TCAATCGACCAATCTCCGATTCATGCAACTGTGCTCTTCTTGCACTCAACATCTGTC 1023
 Db 378 SerIleAspGlnLeuProAspTyMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397
 QY 1024 GATGATACATCGTACGATGTTATGAAGAGAAAGCGCTCAACGTTATACCTTACCTCGG 1083
 Db 398 AspAspThrSerTyAspValMetLysGluLysGlyValAsnValIleProTyLeuArg 417
 QY 1084 CAATCTGGTGTGATTTGGCGGATTAAGTATATGTGTAGAGGACCGTGTCTTACGGCGG 1143
 Db 418 GlnSerTrpValAspLeuAlaAspLysTyMetValGluAlaArgTrpPheTyGlyGly 437
 QY 1144 CACAAACCAAGTTTGAAGAGTATTTGGAGAACTCATCGCAGTTCGATAAGTGGGCCCTGT 1203
 Db 438 HisLysProSerLeuGluGluTyLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457
 QY 1204 ATGTTAACGACATATTTCTCGAGTACAGATTCGTTCAACAAAGGACCGTCGACAGT 1263
 Db 458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477
 QY 1264 TTGTAACAAATACCACTATTTAGTTCGTTGTTGTCATCTCTTCTGCGGCTTGTGATGAT 1323
 Db 478 LeuTyTyLysTyHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497
 QY 1324 TTGGAAACCTCGGTGAAGAGGTGACGAGGAGGTGTGCGGAAATCACTTCAGTGTCTAC 1383
 Db 498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTy 517
 QY 1384 ATGAGTGAATACCACTCGAGGCGGAGGCGCGGAGGAGCGGAGGAGGAGGAGGAGGAG 1443
 Db 518 MetSerAspTyAsnAlaSerGluAlaGluAlaArgLysHisValLysTyLeuIleAla 537
 QY 1444 GAGGTGTGAAGAGTATGATCGGAGGAGGTGTGCGAGGATTCCTCCATTCGCGCAAGAT 1503
 Db 538 GluValTrpLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557
 QY 1504 TTTATAGGATGTCCAGTTCATTTAGGAAGGATGGCGCAGTGTGATGTACCATATGGAGAT 1563
 Db 558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyHisAsnGlyAsp 577
 QY 1564 GGCACGGGCACACACACCTTATTATACATCAACAAATGACCAAGAACCTTATTGAGCCC 1623

Db 578 GlyHisGlyThrGlnHisProIleHisGlnMetThrArgThrLeuPheGluPro 597
Qy 1624 TTTCGA 1629
Db 598 PheAla 599

RESULT 8
US-08-937-540-4
Sequence 4, Application US/08937540
Patent No. 5891697
GENERAL INFORMATION:
APPLICANT: Croteau, Rodney B
APPLICANT: Wise, Mitchell L
APPLICANT: Savage, Thomas J
APPLICANT: Katahira, Eva J
TITLE OF INVENTION: Monoterpene Synthases from Common Sage
TITLE OF INVENTION: (Salvia officinalis)
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESSES:
ADDRESSEE: CHRISTENSEN, O'CONNOR, JOHNSON & KINDNESS
STREET: 1420 FIFTH AVENUE
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA
ZIP: 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/937,540
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Shelton, Dennis K
REGISTRATION NUMBER: 26,997
REFERENCE/DOCKET NUMBER: WSU111254
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206 695 1718
TELEFAX: 206 224 0779
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 591 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-937-540-4

Alignment Scores:
Pred. No.: 1,26e-157 Length: 591
Score: 1633.50 Matches: 305
Percent Similarity: 75.96% Conservative: 109
Best Local Similarity: 55.96% Mismatches: 116
Query Match: 56.17% Indels: 15
Gaps: 7
DB:

US-09-938-956-6 (1-1632) x US-08-937-540-4 (1-591)

Qy 4 AGACGATCCGGAACACTACACCCCTCTCGTTGGGATGTCAACTTCATCCATCGCTTCTC 63
Db 57 ArgArgThrGlyGlyThrGlnProThrLeuTrpAspPheSerThrIleGlnLeuPheAsp 76
Qy 64 AGTGACTATAAGGAGGACCAACACGTCGATTAGGGCTTCTGAGCTGGTCACTTTGGTGAAG 123
Db 77 SerGlnTyrlsGluGluHisLeuMetArgAlaAlaGlyMetIleAlaGlnValasn 96
Qy 124 ATGGAACTGGGAAGAAGAACCGATCAATTCGACAACTTGATGATGATGATGATGATGATG 183
Db 97 MetLeuLeuGlnGluGluValAspSerIleGlnArgLeuGluLeuIleAspLeuArg 116
Qy 184 AGGATGGGGTGTCCGATCATTTCCAAAAGAGTTCAAGAAATCTTGCTCTATATAT 243

Db 117 ArgLeuGlyIleSerCysHisPheAspArgGluIleValGluIleLeuAsnSerIleTyr 136
Qy 244 CTCGACCATCACTATTACAGAACCCCTTTCCAAAAGAGAAAGGATCTCTACTCCACA 303
Db 137 -----TyrThrAsnAsnGluIle-----AspGluSerAspLeuTyrSerThr 150
Qy 304 TCTCTGTGATTTAGGCTCTCAGAGAACATGTTTTCAGTCGCGACAGAGGTATTCGAT 363
Db 151 AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp 170
Qy 364 AGTTTCAAGAACGAGGAGGCT---GAGTTCAAGAAAGCCTTAGCGACGACACACAGAGA 420
Db 171 CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190
Qy 421 TTGTTGCAACTGTATGAAAGCTTCTCTCTGTTGAGCGGAGGCGAAACACCGCTCAGTCA 480
Db 191 LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluThrLeuHisLeu 210
Qy 481 GCGAGGGAATTCGCCACCAAAATTTTGGAGGAAAAGTCAACGAGGCTGTCTGATGCG 540
Db 211 AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu-----ValAspLys 227
Qy 541 GAC-----CTTTTAAACAAGATCGCATATCTTTGGACATCCCTCTTCATTGGAGGATT 594
Db 228 AspIleAsnLeuLeuSerSerIleGluArgAlaLeuGluLeuProThrHisTrpArgVal 247
Qy 595 AAAAGCCCAATGCACCTGTGTGGATCGAATGGTATAGGAAGGCGCGACATGAATCCA 654
Db 248 GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267
Qy 655 GTAGTGTGTGGAGCTTGGCCATCTCGACTTAAATATTTGTTCAAGCACAATTTCAAGAAG 714
Db 268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlu 287
Qy 715 CTCAAAGATCCTTCAGGTGGTGGAGAAATACCTGGTTTGTGAGAAGTGCCTTCGCA 774
Db 288 LeuLysGluAlaSerArgTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307
Qy 775 AGGGATAGACTGTGTGAATGTCTACTTTTGGAACTACTGGGATCATCGAGCCAGTCAGCAT 834
Db 308 ArgAspArgIleValGluCysTyrTrpThrThrGlyValValGluArgGluHis 327
Qy 835 GCAAGTGCAGGATTAATGATGGGAAAGTCAAGCTCTGATTACGGTGTGATCATGATT 894
Db 328 GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrIleAspAspVal 347
Qy 895 TATGATGCTATGGCACCTTAGAAGACTCGAACAAATTCACCTGACCTCATTCGAAAGATGG 954
Db 348 PheAspIleTyrGlyThrLeuGluGluLeuLeuPheThrThrAlaIleGlnArgTrp 367
Qy 955 GATATAAATCAATCGACCAACTTCCGATTACATGCAACTGTGCTTTTCTGCACTCAAC 1014
Db 368 AspIleGluSerMetLysGlnLeuProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387
Qy 1015 AACTTGTGTGATGATACATCGTACGATGTTATGAGGAGAAAGGCGTCAACGTTATACCC 1074
Db 388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407
Qy 1075 TACCTCGGCAATCGTGGGTTGATTGGCGGATAAGTATATGTTAGAGCAGCGTGGTTC 1134
Db 408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTrpTyr 427
Qy 1135 TACGGCGGCGACAAACCAAGTTTGGAGAGTATTGGAGAACTCATGCGCAGTCGATAAGT 1194
Db 428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447
Qy 1195 GGCCCTGTGATTACAGCACATATTTCTCCAGTAACAGATTCGTTCCAAAGAGGAGACC 1254
Db 448 GlyIleProIleLeuSerHisLeuPheArgLeuThrAspSerIleGluGluAsp 467
Qy 1255 GTCGACAGTTGTACAAATACCAAGATTAGTTCGTTGGTTCATCCTTCCTTCGGGCTT 1314

Db 468 AlaGluSerMetHisIleTy-HisAspIleValArgAlaSerCysThrIleLeuArgLeu 487
 QY 1315 GCTGATGATTGGGAACCTCGTGGCAAGAGTGGAGAGGATGCGCAATCACTT 1374
 Db 488 AlaAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProIysSerVal 507
 QY 1375 CAGTCTCATAGTGAATGACTACATGCTCGAGCGGAGCGGAGGAGCAGTGAATGG 1434
 Db 508 GlnCysTyMetAsnGluLysAsnAlaSerGluGluGluAlaArgGluHisValArgSer 527
 QY 1435 CTGATAGCGAGGTGGTGAAGAAGATGATCGAGAGGAGTGGAGGATTCCTCCATC 1494
 Db 528 LeuIleAspGlnThrTrpIysMetMetAsnLysGluMetMetThr-----SerSerPhe 545
 QY 1495 GCGAAGATTTATAGGATGTGCAATGATTAGGAAGTGGCGCAGTGTGATGATC--- 1551
 Db 546 SerIleTyPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyGln 565
 QY 1552 CATATGAGATGGGACCGGCACACACACCTATTATACATCAACAAATGACCGAAC 1611
 Db 566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585
 QY 1612 TTATTCGAGCCCTTT 1626
 Db 586 LeuPheAspArgTy 590

RESULT 9

US-09-398-395A-24
 ; Sequence 24, Application US/09398395A
 ; Patent No. 6468772
 ; GENERAL INFORMATION:
 ; APPLICANT: Chappell, Joseph
 ; APPLICANT: No. 64687721, Joseph P.
 ; APPLICANT: Starks, Courtney M.
 ; APPLICANT: Manna, Kathleen R.
 ; TITLE OF INVENTION: SYNTHASE
 ; FILE REFERENCE: 07678-025001
 ; CURRENT APPLICATION NUMBER: US/09/398,395A
 ; CURRENT FILING DATE: 1999-09-17
 ; PRIOR APPLICATION NUMBER: 60/100,993
 ; PRIOR FILING DATE: 1998-09-18
 ; PRIOR APPLICATION NUMBER: 60/130,628
 ; PRIOR FILING DATE: 1999-04-22
 ; PRIOR APPLICATION NUMBER: 60/150,262
 ; PRIOR FILING DATE: 1999-08-23
 ; NUMBER OF SEQ ID NOS: 58
 ; SEQ ID NO 24
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; LENGTH: 591
 ; TYPE: PRT
 ; ORGANISM: Salvia officinalis
 US-09-398-395A-24

Alignment Scores:
 Pred. No.: 1,26e-157 Length: 591
 Score: 1633.50 Matches: 305
 Percent Similarity: 75.96% Conservative: 109
 Best Local Similarity: 55.96% Mismatches: 116
 Query Match: 56.17% Indels: 15
 DB: 4 Gaps: 7

US-09-938-956-6 (1-1632) x US-09-398-395A-24 (1-591)

QY 4 AGACGATCCGGAACACTACACCTTCTCGTGGGATGTCACATTCATCCAAATCGCTTC 63
 Db 57 ArgAspThrGlyGlyTy-GlnProThrLeuTrpAspPheSerThrIleGlnLeuPheAsp 76
 QY 64 AGTGACTATAGGAGCAACACGATGATTAGGCTTCTGAGCTGGTGCATTTGGTGAAG 123
 Db 77 SerGluTyLysGluGluLysHisLeuMetCargAlaGlyMetIleAlaGlnValAsn 96
 QY 124 ATGGAACTGGGAAGAAACGATCAAAATTCGCAACTTGATGATCGATGCTGCGAG 183

Db 97 MetLeuLeuGlnGluGluValAspSerIleGlnArgLeuGluLeuIleAspSerLeuArg 116
 QY 184 AGGATGGGCTGTCGCGATCATTTCCAAATGATGTTCAAAGAAATCTGTCCTTATATAT 243
 Db 117 ArgLeuGlyIleSerCysHisPheAspArgGluIleValGluIleLeuAsnSerTyLys 136
 QY 244 CTCGACCATCACTATTACAAGAACCTTTTCCAAAAGAAAGAGGATCTCTACTCCACA 303
 Db 137 -----TyThrAsnAsnGluIle-----AspGluSerAspLeuTySerThr 150
 QY 304 TCTCTTCGATTTAGCTTCCTCAGAGAACATGGTTTCAAGTCCCAACAGAGGATTCGAT 363
 Db 151 AlaLeuArgPheLysLeuLeuArgGlnTyAspPheSerValSerGlnGluValPheAsp 170
 QY 364 AGTTTCAAGAACAGAGGAGGCT---GAGTTCAAAAGAAAGCTTAGCCACACACACAGAGA 420
 Db 171 CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190
 QY 421 TTGTTGCAACTGTATGAGCTTCTTCTGTTCCGAGAGGCGAAACACCGCTCGAGTCA 480
 Db 191 LeuLeuGlnLeuTyGluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu 210
 QY 481 GCGAGGAAATTCGCCCAAAATTTTGGAGGAAAGTGAACGAGGCTGTTGTTGATGCG 540
 Db 211 AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu-----ValAspLys 227
 QY 541 GAC-----CTTTTAAACAAGATCGCATATTCTTTGACATCCCTCTTCTCATTCGAGATT 594
 Db 228 AspIleAsnLeuLeuSerIleGluAlaLeuGluLeuProThrHisTrpArgVal 247
 QY 595 AAAAGGCCAAATGCACCTGTGTGATCGAATGGTATAGGAAGAGGCCGACATGATCCA 654
 Db 248 GlnMetProAsnAlaArgSerPheIleAspAlaTyLysArgArgProAspMetAsnPro 267
 QY 655 GTAGTGTGAGCTGCGCATCTGCACTTAATATTGTTCAAGCAAAATTCAGAAAGAG 714
 Db 268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlu 287
 QY 715 CTCAAAGAATCCTTCAGGTGGTGGAGAAATACCTGGGTTTGTTCAGAAAGCTGCCCTTCGA 774
 Db 288 LeuLysGluAlaSerArgTrpTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307
 QY 775 AGGATAGCTGTTGGATGCTACTTTTGGAACTACCTGGGATCGGATCATCGAGCCAGCAT 834
 Db 308 ArgAspArgIleValGluCysTyTrpThrThrGlyValValGluArgGluHis 327
 QY 835 GCAAGTCAGAGTAATGATGGCAAGTCAACGCTCTGATTACGGTGTGATGATATT 894
 Db 328 GlyTyGluArgIleMetLeuThrLysIleAsnAlaLeuValThrThrIleAspVal 347
 QY 895 TATGATCTATGACCTTTAGAGAACTCGAACAAATTCACCTGACCTCATTCGAGATCG 954
 Db 348 PheAspIleTyGlyThrLeuGluGluLeuGlnLeuPheThrThrAlaIleGlnArgTrp 367
 QY 955 GATATAAATCAATTCGACCAACTTCCGATTAATGCAACTGCTGCTTTCTTGCACTCAAC 1014
 Db 368 AspIleGluSerMetLysGlnLeuProProTyMetGlnIleCysTyLeuAlaLeuPhe 387
 QY 1015 AACTTCGTCGATGATACATCGTACGATGTTATGAGAGAGAAAGGCTCAACGTTATACC 1074
 Db 388 AsnPheValAsnGluMetAlaTyAspThrLeuArgAspLysGlyPheAsnSerThrPro 407
 QY 1075 TACTCGGCAATCGTGGGTTGATTGGCGGATTAAGTATATGATAGAGCAACGGTGGTTC 1134
 Db 408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyLeuIleGluAlaLysTrpTy 427
 QY 1135 TACGGCGGCACAAAACAGTTTGGAGAGTATTTGGAGAACTTCAGAGCAGTCGATAAGT 1194
 Db 428 TyrMetGlyHisLysProSerLeuGluGluTyMetLysAsnSerTrpIleSerIleGly 447
 QY 1195 GGGCCCTGTATGTTAAACGACATATTCTTCCGAGTAAACAGATTCGTTCAACAGAGGACC 1254
 Db 448 GlyIleProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluAsp 467

QY 1135 TAGCGCGGCAAAACCAAGTTTGGAGAGTATTGGAGAACTCATGGCAGTCGATAAGT 1194
Db 428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447
QY 1195 GGGCCCTGTATGTTAAACGACATATCTCCGAGTACAGATTGGTTCACAAAGGAGACC 1254
Db 448 GlyLeuProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluGluAsp 467
QY 1255 GTCGACAGTTGTACAAATACCAAGATTAGTTGCTGTCATCCCTCGTTCGCGCTT 1314
Db 468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487
QY 1315 GCTGATCATTTGGAACTCGGTGGAAGAGTGCAGAGGGGATGCGCGAATCACTT 1374
Db 488 AlaAspAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProLysSerVal 507
QY 1375 CAGTGTACATGAGTCACTACATCGAGCGGAGCGGAGCGGAGCAACGTAATGG 1434
Db 508 GlnCysTyrMetAsnGluLysAsrAlaSerGluGluAlaArgGluHisValArgSer 527
QY 1435 CTGATAGCGAGGTGTGGAAGACATGAATCCGAGAGGGTGTGGAAGATTCTCCATTC 1494
Db 528 LeuIleAspGlnThrTrpLysMetMetAsnLysGluMetMetThr-----SerSerPhe 545
QY 1495 GCGAAGATTTTATAGCATGTCAGTTGATTAGGAAGGATGGCGAGTGTGATGTAC--- 1551
Db 546 SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln 565
QY 1552 CATATGAGATGGGACGGCACACACCTATTATATCATCAACAAATGACAGAAC 1611
Db 566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585
QY 1612 TTATTCGAGCCCTTT 1626
Db 586 LeuPheAspArgTyr 590

RESULT 12

US-09-903-012B-24
; Sequence 24, Application US/09903012B
; Patent No. 6569656
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 65696561, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012B
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Salvia officinalis
US-09-903-012B-24

Alignment Scores:
Pred. No.: 1,26e-157 Length: 591
Score: 1633.50 Matches: 305
Percent Similarity: 75.96% Conservative: 109
Best Local Similarity: 55.96% Mismatches: 116
Query Match: 56.17% Indels: 15
DB: 4 Gaps: 7

US-09-938-956-6 (1-1632) x US-09-903-012B-24 (1-591)

QY 4 AGACGATCCGGAACCTACACACCTTCTCGTTGGGATGTCACCTTCATCCATCGCTTCTC 63
Db 57 ArgArgThrGlyGlyTyrGlnProThrLeuTrpAspPheSerThrIleGlnLeuPheAsp 76
QY 64 AGTGACTATAAGGAGGACAAACACGTGATTAGGCTTCTGAGCTGGTGCCTTCTGCTGAAG 123
Db 77 SerGluTyrLysGluGluLysHisLeuMetArgAlaAlaGlyMetIleAlaGlnValAsn 96
QY 124 ATGGAACCTGGAGAGAAAGAACCGATCAAAATTCGACAACTTGATGTCATCGATCGACTG 183
Db 97 MetLeuLeuGlnGluGluValAspSerIleGlnArgLeuGluLeuIleAspLeuArg 116
QY 184 AGGATGGGCTGTGCGATCATTTCCAAAATGAGTTTCAAGAAATCTTCTCTCTATATAT 243
Db 117 ArgLeuGlyIleSerCysHisPheAspArgGluIleValGluIleLeuAsnSerLysTyr 136
QY 244 CTCGACCATCATATTATCAAGAACCCCTTTCCAAAGAAAGAAAGGATCTCTACTCCACA 303
Db 137 -----TyrThrAsnAsnGluIle-----AspGluSerAspLeuTyrSerThr 150
QY 304 TCTCTTGCATTTAGGCTCTCTCAGAGAACATGTTTTTCAAGTCGCAACAAGAGGTATTCGAT 363
Db 151 AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp 170
QY 364 AGTTTCAAGAACGAGGAGGGT---GAGTTCAAGAAAGCCCTTAGCGACGACACACGAGGA 420
Db 171 CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190
QY 421 TTGTTGCACTGTATGAAGCTTCTTCTGTTGCGGAGGCGGAAACCCAGCTCCAGATCA 480
Db 191 LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu 210
QY 481 GCGAGGGAATCCGCCACCAAAATTTTGGAGGAAAAAGTGAACGAGGGTGTGTGATGCGC 540
Db 211 AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu-----ValAspLys 227
QY 541 GAC-----CTTTTAAACAAGATCGCATATCTTTGGACATCCCTTCTCATTTGGAGGATT 594
Db 228 AspIleAsnLeuLeuSerSerIleGluArgAlaLeuGluLeuProThrHisTrpArgVal 247
QY 595 AAAAGCCAAATGCACTGTGTGGATCGAATGGTATAGGAAGAGCCGCGACATGAATCCA 654
Db 248 GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267
QY 655 GTAGTGTGGAGCTTGCCCATCTCGACTTAATATTTGTCAAGCAACATTTCAAGAGAG 714
Db 268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlnGlu 287
QY 715 CTCAAAGATCCCTCAGGTGGTGGAGAAATACATGGGTTTGTGAGAAGCTGCCCTTCGCA 774
Db 288 LeuLysGluAlaSerArgTrpTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307
QY 775 AGGGATAGACTGGTGGAAATGCTACTTTTGGATATCTGGGATCATCGAGCCAGTCAGCAT 834
Db 308 ArgAspArgIleValGluCysTyrTyrTrpThrGlyValValGluArgGluHis 327
QY 835 GCAGTGCAGGATATGATGGGCAAGTCACGCTCTGATTACCGTGTATCGATCATATT 894
Db 328 GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrIleAspAspVal 347
QY 895 TATGATGCTATGGCACCTTAGAAGAACTCGAACAATTCACCTGACCTCATTCGAAAGATGG 954
Db 348 PheAspIleTyrGlyThrLeuGluGluLeuGlnLeuPheThrThrAlaIleGlnArgTrp 367
QY 955 GATATAAATCAATCGACCAACTTCCGATTCATGCAACTGTGCTTCTTCTGCACTCAAC 1014
Db 368 AspIleGluSerMetLysGlnLeuProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387
QY 1015 AACTTCGTGATGATACATCGTACGATGTTATGAGGAGAGAAAGGGGTCAACGTATACCC 1074
Db 388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407

1075 TACCTGCGGCAATCGTGGTGGATTGGCGGATAGTATATGTTAGAGGACGCGTGGTTC 1134
Db TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTrpTyr 427
1135 TAGCGCGGCGCAAAACCAAGTTTGAAGATTTTGGAGAACTCATGCGATCGATAGT 1194
Db TyrMetGlyHisLysProSerLeuGluGlyTyrMetLysAsnSerTrpIleSerIleGly 447
1195 GGGCCCTGTATGTTAAACGACATATCTTCGAGTAACAGATTCGTTCCACAAAGGACG 1254
Db GlyIleProIleLeuSerHisLeuPheArgLeuThrAspSerIleGluGluGluAsp 467
1255 GTCGACAGTTGTACAAATACACAGTTAGTTGCTTGGTTCATCTTCGCGCTT 1314
Db AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487
1315 GCTGATGATTTGGGACCTCGTGAAGAGTGTGACGAGGGGATGTCGCGAAATCACTT 1374
Db AlaAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProLysSerVal 507
1375 CAGTCTACATGAGTACCTACAAATGATCGAGGCGGCGGAGGACGCTGAAATGG 1434
Db GlnCysTyrMetAsnGluLysAsnAlaSerGluGluGluAlaArgGluHisValArgSer 527
1435 CTGATAGCGAGGTGTGAAGAGATGAATCGCGAGAGGCTGTGAGGATTCCTCATTC 1494
Db LeuIleAspGlnThrTrpLysMetMetAsnLysGluMetMetThr-----SerSerPhe 545
1495 GCGAAGAGTTTATAGTATGTCAGTTCGATTTAGGAAGGATGGCGAGTTCATGATAC--- 1551
Db SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln 565
1552 CATATGAGATGGGACGCGACACACACCTTATATATACATCAAAATGACAGAAC 1611
Db HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585
1612 TTATTCGAGCCCTTT 1626
Db LeuPheAspArgTyr 590

RESULT 13
US-09-900-797-24

Sequence 24, Application US/09900797
Patent No. 6645762

GENERAL INFORMATION:

APPLICANT: Chappell, Joseph
APPLICANT: No. 66457621, Joseph P.

APPLICANT: Starks, Courtney M.

APPLICANT: Manna, Kathleen R.

TITLE OF INVENTION: SYNTHASES

FILE REFERENCE: 07678-025001

CURRENT APPLICATION NUMBER: US/09/900,797

PRIOR FILING DATE: 2001-07-06

PRIOR APPLICATION NUMBER: US/09/398,395

PRIOR FILING DATE: 1999-09-17

PRIOR APPLICATION NUMBER: 60/130,628

PRIOR FILING DATE: 1999-04-22

PRIOR APPLICATION NUMBER: 60/150,262

PRIOR FILING DATE: 1999-08-23

NUMBER OF SEQ ID NOS: 58

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 24

LENGTH: 591

TYPE: PRT

ORGANISM: Salvia officinalis

US-09-900-797-24

Alignment Scores:

Pred. No.: 1,266-157

Score: 1633.50

Percent Similarity: 75.96%

Best Local Similarity: 55.96%

Query Match: 15

Length: 591

Matches: 305

Conservative: 109

Mismatches: 116

Indels: 15

DB: 4 Gaps: 7
US-09-938-956-6 (1-1632) x US-09-900-797-24 (1-591)
QY 4 AGACGATCGGAAACTAGAACCTTCTCGTGGGATGTCAACTTCATCCATCGCTTC 63
Db ArGArgThrGlyGlyTyrGlnProThrLeuTrpAspPheSerThrIleGlnLeuPheAsp 76
QY 64 AGTCACTATAAGGAGACAAACACGTTGATTAGGGCTTCTGAGCTGGTCACTTTGTGAAG 123
Db SerGluTyrLysGluGluLysHisLeuMetArgAlaAlaGlyMetIleAlaGlnValAsn 96
QY 124 ATCGAACTCGAAGAAACCGATCAAAATTCACAACTTGAATTCATCGATCGATTGAC 183
Db MetLeuLeuGlnGluGluValAspSerIleGlnArgLeuGluLeuIleAspLeuArg 116
QY 184 AGGATGGGCTGTCCGATCATTTCCAAAATGAGTTCAAAAGAAATCTTGTCTCTATATAT 243
Db ArgLeuGlyLysCysHisPheAspArgGluIleValGluIleLeuAsnSerLysTyr 136
QY 244 CTCGACCATCATTTATCAAGAACCCCTTTCCAAAAGAAAGAAAGGATCTTACTCCACA 303
Db TyrThrAsnAsnGluIle-----AspGluSerAspLeuTyrSerThr 150
QY 304 TCTCTTTCATTTAGGCTCTCAGAGAACATGTTTCAAGTCGACACAGAGGATTTCGAT 363
Db AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp 170
QY 364 AGTTTCAAGAACGAGGAGGGT---GAGTTCAAGAAAGCCTTTAGCGACGACACAGAGA 420
Db CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190
QY 421 TTGTTCCAACTGTATCAAGCTTCTTCTGTCGAGGAGGCGAAACACGCTCGAGTCA 480
Db LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluThrLeuHisLeu 210
QY 481 GCGAGGAAATTCGCCACCAAAATTTTGGAGGAAAGAGTGAACGAGGGTGTGTGTATGCG 540
Db AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu-----ValAspLys 227
QY 541 GAC-----CTTTTACAAGATCGCATATTTCTTGGACATCCCTCTTCATTTGGAGATT 594
Db AspIleAsnLeuLeuSerSerIleGluArgAlaLeuGluLeuProThrHisTrpArgVal 247
QY 595 AAAAGGCCAAATCGACTGTGTGATCGAATGATAGGAAGAGGCCGACATCAATCCA 654
Db GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267
QY 655 GTAGTGTGGAGCTTCCCATCTCTGACTTAAATATTTTCAAGACCAATTTCAAGAGAG 714
Db ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlnGlu 287
QY 715 CTCAAAAGAAATCCTTCAGGTGGTGGAGAAATACCTGCGGTTTGTGAGAGCTGCCCTTCGCA 774
Db LeuLysGluAlaSerArgTrpTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307
QY 775 AGGATAGACTGCTGGAATCTACTTTTGAATCTGGAATCTATCGGATCATCGAGCCAGCAT 834
Db ArgAspArgIleValGluCysTyrTyrTrpThrGlyValValGluArgArgGluHis 327
QY 835 GCAAGTCGAGGATATGATGGCGCAAGTCACGCTCTGATTCAGGTGATCGATGATATT 894
Db GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrThrIleAspAspVal 347
QY 895 TATGATGCTATGCGACCTTTAGGAAGAACTCGAACAAATTCAGTCACTTCATTCGAAATGG 954
Db PheAspIleTyrGlyThrLeuGluGluLeuLeuPheThrThrAlaIleGlnArgTrp 367
QY 955 GATATAACTCAATCGACCACTTCGCGATACATCACTGCTGCTTCTTTCGACCTCAAC 1014
Db AspIleGluSerMetLysGlnLeuProProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387
QY 1015 AACTTCGTCGATGATACATCGTACGATGTTTATGAGGAGAAAGCGCTCAACGTTATACCC 1074

Db 388 AsnPheValAsnGluMetAlaTyAspThrLeuArgAspLysGlyPheAsnSerThrPro 407
Qy 1075 TACCTGGCGCAATCGTGGTGGTATTTGGCGGATAAGTATATGTAGAGGACAGGTGGTTC 1134
Db 408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuLeuGluAlaLysTrpTyr 427
Qy 1135 TAGCGCGGACAAACCAAGTTTGAAGAGATTTGGAGAACTCATGCGAGTCGTAAGT 1194
Db 428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpLeuSerIleGly 447
Qy 1195 GGGCCCTGTATGTTAAACGACATATCTTCGAGTAAACAGATTCGTTTCACAAAGAGACC 1254
Db 448 GlyLeuProIleLeuSerHisLeuPhePheArgLeuThrAspSerIleGluGluAsp 467
Qy 1255 GTCGACAGTTTGTACAAATACACAGATTTAGTTGGTTCATCCTTGGTTCGGGCTT 1314
Db 468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487
Qy 1315 GCTCATGATTTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTCGCGAATCATT 1374
Db 488 AlaAspMetGlyThrSerLeuAspLeuValGluArgGlyAspValProLysSerVal 507
Qy 1375 CAGTCTACATGACTACTAATGCAATCGAGCGGAGCGGCGGAGACGAGTGAATGG 1434
Db 508 GlnCysTyrMetAsnGluLysAsnAlaSerGluGluAlaArgGluHisValArgSer 527
Qy 1435 CTGATAGCGGAGGTGTGGAAGAAGATGAATGCGGAGAGGTGTGCAAGGATTTCCATTC 1494
Db 528 LeuIleAspGlnThrTrpLysMetMetAsnLysGluMetMetThr-----SerSerPhe 545
Qy 1495 GGCAAGATTTTATAGGATGTGCAGTTGATTAGCAAGGATGGCGAGTGTGATGATC--- 1551
Db 546 SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln 565
Qy 1552 CATAATGAGATGGCGACGACCAACACACCTATTATATCATCAACAAATGACCGAACC 1611
Db 566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585
Qy 1612 TTATTCGAGCCCTT 1626
Db 586 LeuPheAspArgTyr 590

RESULT 14

US-08-937-540-6
Sequence 6, Application US/08937540
Patent No. 5891697
GENERAL INFORMATION:
APPLICANT: Croteau, Rodney B
APPLICANT: Wise, Mitchell L
APPLICANT: Savage, Thomas J
APPLICANT: Katahira, Eva J
TITLE OF INVENTION: Monoterpene Synthases from Common Sage
TITLE OF INVENTION: (Salvia officinalis)
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESS: CHRISTENSEN, O'CONNOR, JOHNSON & KINDNESS
STREET: 1420 FIFTH AVENUE
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA
ZIP: 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/937,540
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Shelton, Dennis K

REGISTRATION NUMBER: 26,997
REFERENCE/DOCKET NUMBER: WSUR111254
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206 695 1718
TELEFAX: 206 224 0779
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 589 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-937-540-6
Alignment Scores:
Pred. No.: 1,82e-149 Length: 589
Score: 1553.50 Matches: 291
Percent Similarity: 72.48% Conservative: 104
Best Local Similarity: 53.39% Mismatches: 141
Query Match: 53.42% Indels: 9
DB: 2 Gaps: 3
US-09-938-956-6 (1-1632) x US-08-937-540-6 (1-589)
Qy 1 ATGAGACGATCGGAAACTACAACTTCGCTGGGATGTCACTTCATCAATCGCTT 60
Db 51 IleAArgSerGlyAspTyrGlnProSerLeuTrpAspPheAsnTyrIleGlnSerLeu 70
Qy 61 CTCAGTGCATATAAGGAGGACAAACACGCTGATTAGGGCTTCGAGCTGGTCACTTTGGTG 120
Db 71 AsnThrProTyrLysGluGlnArgHisPheAsnArgGlnAlaGluLeuIleMetGlnVal 90
Qy 121 AAGATGGAACTCGGAGAGAAACCGGATCAAAATTCGCAACTTGAAGTTGATGATGACTTG 180
Db 91 ArgMetLeuLeuLysValLysMetGluAlaIleGlnGlnLeuGluLeuIleAspLeu 110
Qy 181 CAGAGATGGGCTCTCCGATCATTTCCAAATAGTTTCAAGAAATCTGTCTCTATA 240
Db 111 GlnTyrLeuGlyLeuSerTyrPhePheGlnAspGluIleLysGlnIleLeuSerIle 130
Qy 241 TATCTCGACCATCATATTACAAAGAACCTTTTCCAAAGAGAAAGGATCTCTACTCC 300
Db 131 HisAsnGluProArgTyrPheHisAsn-----AsnAspLeuTyrPhe 144
Qy 301 ACATCTTTGCAATTTAGGCTCTCAGAGAAACATGTTTTCAGTGCACAAAGAGTATTC 360
Db 145 ThrAlaLeuGlyPheArgIleLeuArgGlnHisGlyPheAsnValSerGluAspValPhe 164
Qy 361 GATAGTTTCAAGAACGAGGAG---GCTGAGTTTCAAGAAAGCTTAGCCGACGACACAGA 417
Db 165 AspCysPheLysIleGluLysCysSerAspPheAsnAlaAsnLeuAlaGlnAspThrLys 184
Qy 418 GGATTTTCCAACTGTATGAAGCTTCCTTTCTGTTGACGGAAGGGGAAACCGCTCGAG 477
Db 185 GlyMetLeuGlnLeuTyrGluAlaSerPheLeuLeuArgGluGlyGluAspThrLeuGlu 204
Qy 478 TCAGCGAGGAATTCGCCACCAATTTTGGAGGAAAGTGAACGAGGCTGT----- 531
Db 205 LeuAlaArgPheSerThrArgSerLeuArgGluLysPheAspIleGlyAspGlu 224
Qy 532 GTTGATGGCGACCTTTTAAACAAGAATCGCATATTTCTTGGACATCCCTCTTCATGGAGG 591
Db 225 IleAspGluAspLeuSerSerTrpIleArgHisSerLeuAspLeuProLeuHisTrpArg 244
Qy 592 ATTAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAGAGGCCCGACATGAAT 651
Db 245 ValGlnGlyLeuGluAlaArgTrpPheLeuAspAlaTyrAlaArgArgProAspMetAsn 264
Qy 652 CCAGTAGTGTGGAGCTTCGCATATCTCGACTTAAATATTGTTCAAGCACAATTTCAAGAA 711
Db 265 ProLeuIlePheLysLeuAlaLysLeuAsnIleValGlnAlaThrTyrGlnGlu 284
Qy 712 GAGCTCAAGAATCTTTCAGGTGGTGGAATAACTGGGTTTGTTCAGAGGCTGCCCTTC 771

```
Db 285 GluLeuLysAspIleSerArgTrpTrpAsnSerSerCysLeuAlaGluLysLeuProPhe 304
QY 772 GCAGGATAGACTGTGAATGCTACTTTTGGATATCTGGGATCATCGACCGCTGAG 831
Db 305 ValArgAspArgIleValGluCysPhePheTrpAlaIleAlaAlaPheGluProHisGln 324
QY 832 CATCAAGTCCAGGATAATGATGGCAAGCAACGCTCTGATACCGTGTATCATCAT 891
Db 325 TyrSerTyrGlnArgLysMetAlaAlaValIleIleThrPheIleThrIleIleAsp 344
QY 892 ATTATGATGTCTATGGCACTTAGAGAACTCGAACAATTCACGACCTCATTCGAAGA 951
Db 345 ValTyrAspValTyrGlyThrIleGluLeuGluLeuLeuLeuLeuLeuLeuLeuLeu 364
QY 952 TGGGATATAAATCAATCAACGACCACTTCCCGATTACATCAACTGTGCTTCTTCAC 1011
Db 365 TrpAspAsnLysSerIleSerGlnLeuProTyrTyrMetGlnValCysTyrLeuAlaLeu 384
QY 1012 AACAACTTCGTGATGATACATGCTACGATGTTGGAAGAGTATTTGGAGAACTCATG 1071
Db 385 TyrAsnPheValSerGluArgAlaTyrAspIleLeuLysAspGlnHisPheAsnSerIle 404
QY 1072 CCTACTCTGGCAACCTGCTGGTGTATTTGGCGGATAGTATATGTTAGGACGCGG 1131
Db 405 ProTyrLeuGlnArgSerTrpValSerLeuValGluGlyTyrLeuLysGluAlaTyrTrp 424
QY 1132 TTCTACGGCGGCGCACAAACCAATTTGGAAGAGTATTTGGAGAACTCATGCGAGTCGATA 1191
Db 425 TyrTyrAsnGlyTyrLysProSerLeuGluGlyTyrLeuAsnAsnAlaLysIleSerIle 444
QY 1192 AGTGGCCCTGTATGTTAAGCAGCATATTTCCGAGTAACAGATTCGTTTCACAAAGGAG 1251
Db 445 SerAlaProThrIleIleSerGlnLeuTyrPheThrLeuAlaAsnSerIleAspGluThr 464
QY 1252 ACCGTGCGAGTTGTACAAATACCAACGATTTAGTTGGTGTCTCTCTCTCTCTCTGGG 1311
Db 465 AlaIleGluSerLeuTyrGlnTyrHisAsnIleLeuTyrLeuSerGlyThrIleLeuArg 484
QY 1312 CTTGCTGATGATTGGGAACCTCGTGGAGAGGTGAGCAGAGGGAGTGTCCGGAATCA 1371
Db 485 LeuAlaAspAspLeuGlyThrSerGlnHisGluLeuGluArgGlyAspValProLysAla 504
QY 1372 CTTCACTGTACATGATGACTACATGATCGAGCGGCGGCGGAGCAACGCTGAAA 1431
Db 505 IleGlnCysTyrMetAsnAspThrAsnAlaSerGluArgGluAlaValGlnHisValLys 524
QY 1432 TGCGCTGATACGGAGGTGTGGAAGAAGATGAATGCGGAGAGGTGTGGAAGGATTCCTCA 1491
Db 525 PheLeuIleArgGluAlaTrpLysGluMetAsnThrValThrThrAlaSerAspCysPro 544
QY 1492 TTCGGCAAGATTTTATAGATGTGCAGTGTGATTTAGGAGGATGCGCGAGTTGATGTAC 1551
Db 545 PheThrAspAspLeuValAlaAlaAlaAsnLeuAlaArgAlaAlaGlnPheIleTyr 564
QY 1552 CATATGAGAGTGGCGACGCGCACACACCCCTATTATATACATCAACAATGACCAAGC 1611
Db 565 LeuAspGlyAspGlyHisGlyValGlnHisSerGluIleHisGlnGlnMetGlyGlyLeu 584
QY 1612 TTATTCAGCCCTTT 1626
Db 585 LeuPheGlnProTyr 589
```

RESULT 15

```
US-09-398-395A-54
; Sequence 54, Application US/09398395A
; Patent No. 6468772
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 6468772, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
```

```
; CURRENT APPLICATION NUMBER: US/09/398,395A
; CURRENT FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 54
; LENGTH: 590
; TYPE: PRT
; ORGANISM: Salvia officinalis
; US-09-398-395A-54

Alignment Scores:
Pred. No.: 1,82e-149 Length: 590
Score: 1533.50 Matches: 291
Percent Similarity: 72.48% Conservative: 104
Best Local Similarity: 53.39% Mismatches: 141
Query Match: 53.42% Indels: 9
DB: 4 Gaps: 3

US-09-938-956-6 (1-1632) x US-09-398-395A-54 (1-590)
QY 1 ATGAGACGATCCGAAACTACAACCTTCTGTTGGATGTCACATTCATCCAACTGCTT 60
Db 51 IleArgArgSerGlyAspTyrGlnProSerLeuTrpAspPheAsnTyrIleGlnSerLeu 70
QY 61 CTCAGTCACTATAGGAGGACAAACACGTCGATTAGGCTTCTGAGCTGGTCACTTTGGTG 120
Db 71 AsnThrProTyrLysGluGlnArgHisPheAsnArgGlnAlaGluLeuIleMetGlnVal 90
QY 121 AAGATGGAACCTGGAGAAAGAACGGATCAAAATTCGAACACTTCAGTTGATGATGACTTG 180
Db 91 ArgMetLeuLeuLysValLysMetGluAlaIleGlnGlnLeuGluLeuIleAspLeu 110
QY 181 CAGAGGATGGGCTGTCGATCATTTCCAAAATGAGTTCAAGAAATCTTGTCTCTATA 240
Db 111 GlnTyrLeuGlyLeuSerTyrPhePheGlnAspGluIleLysGlnIleLeuSerSerIle 130
QY 241 TATCTCAGCATCACATATTACAAAGACCTTTTCCAAAAGAAAGAGGATCTCTACTCC 300
Db 131 HisAsnGluProArgTyrPheHisAsn-----AsnAspLeuTyrPhe 144
QY 301 ACATCTTCGATTTAGCTCTCAGAGAACATGTTTCAAGTCGCACAGAGGATTTTC 360
Db 145 ThrAlaLeuGlyPheArgIleLeuArgGlnHisGlyPheAsnValSerGluAspValPhe 164
QY 361 GATAGTTTCAAGAACGAGGAG--GGTGAGTTCAAGAAAGCCTTAGCGACGACACAG 417
Db 165 AspCysPheLysIleGluLysCysSerAspPheAsnAlaAsnLeuAlaGlnAspThrLys 184
QY 418 GGATTTGTCAACTGTATGAAGCTTCCTTTCTGTGAGGAGGCGAAACACAGCTCGAG 477
Db 185 GlyMetLeuGlnLeuTyrGluAlaSerPheLeuLeuArgGluGlyGluAspThrLeuGlu 204
QY 478 TCAGCGAGGGAATTCGCCACCAATTTTTCGAGGAAAGAGTCAACGAGGCTGGT----- 531
Db 205 LeuAlaArgArgPheSerThrArgSerLeuArgGluLysPheAspGluGlyAspGlu 224
QY 532 GTTGATGCGGACCTTTTAAACAAGATCGCATATTTTGTGACATCCCTCTTCATTCAGG 591
Db 225 IleAspGluAspLeuSerSerTrpIleArgHisSerLeuAspLeuProLeuHisTrpArg 244
QY 592 ATTAAAGGCCAATGCACTGCTGTGATCGATCGATGATAGGAGAGGCGCCGACATCAAT 651
Db 245 ValGlnGlyLeuGluAlaArgTrpPheLeuAspAlaTyrAlaArgArgProAspMetAsn 264
QY 652 CCAGTAGTGTGGAGCTTGCCACTACTGACTTAAATATTTTCAAGCACAATTTCAAGAA 711
Db 265 ProLeuIlePheLysLeuAlaLysLeuAsnPheAsnIleValGlnAlaThrTyrGlnGlu 284
```



```

QY 712 GAGCTCAAGAAATCCTCAGGTGGTGGAGAAATACTGGTTTGTGTGAGAACTGCCCTTC 771
Db      |||
QY 285 GluLeuLysAspIleSerArgTrpAsnSerSerCysLeuAlaGluLysLeuProPhe 304
Db      |||
QY 772 GCAAGGATAGACTGGTGGATGCTACTTTTGGAAATCTGGGATCATCGACCGACGTCAG 831
Db      |||
QY 305 ValArgAspArgIleValGluCysPheTrpAlaIleAlaAlaPheGluProHisGln 324
Db      |||
QY 832 CATGCAAGTGCAGGATATGATGGGCAAGTCAACGCTCTGATTACGGTGCATCGATGAT 891
Db      |||
QY 325 TyrSerTyrGlnArgLysMetAlaAlaValIleIleThrPheIleThrIleIleAspAsp 344
Db      |||
QY 892 ATTATGATGCTATGACCACTTAGAAGAACTCGAACAATTCACCTGACCTCATTCGAAGA 951
Db      |||
QY 345 ValTyrAspValTyrGlyThrIleGluGluLeuLeuLeuLeuThrAspMetIleArgArg 364
Db      |||
QY 952 TGGGATATAACTCAATCAGCAACTCCCGATTACATGCAACTGTGCTTTCTTGTGCACTC 1011
Db      |||
QY 365 TrpAspAsnLysSerIleSerGlnLeuProTyrTyrMetGlnValCysTyrLeuAlaLeu 384
Db      |||
QY 1012 AACAACTTCGTGATGATACATCGTACGATGTTATGAAGAGAAAGCGTCAACGTTATA 1071
Db      |||
QY 385 TyrAsnPheValSerGluArgAlaTyrAspIleLeuLysAspGlnHisPheAsnSerIle 404
Db      |||
QY 1072 CCTACCTCGCGCAATCCTGGTTCGATTGCGGATTAAGTATATGCTAGAGCGACGGTGG 1131
Db      |||
QY 405 ProTyrLeuGlnArgSerTrpValSerLeuValGluGlyTyrLeuLysGluAlaTyrTrp 424
Db      |||
QY 1132 TTCTACGGCGGCGCAAAACCAAGTTTGAAGAGTATTTGGAGAACTCATGGCAGTCGATA 1191
Db      |||
QY 425 TyrTyrAsnGlyTyrLysProSerLeuGluGluTyrLeuAsnAsnAlaLysIleSerIle 444
Db      |||
QY 1192 AGTGGCCCTGTATGTTAAGCAGCATATTTCTCCAGTAAACAGATTCGTTCCAAAGAG 1251
Db      |||
QY 445 SerAlaProThrIleIleSerGlnLeuTyrPheThrLeuAlaAsnSerIleAspGluThr 464
Db      |||
QY 1252 ACCGTCGACAGTTCGATCAATACCACTTAGTTAGTTCTGTCATCCTTCTGTCGGG 1311
Db      |||
QY 465 AlaIleGluSerLeuTyrGlnTyrHisAsnIleLeuTyrLeuSerGlyThrIleLeuArg 484
Db      |||
QY 1312 CTTGCTGATGATTTGGGAACCTCGGTGGAAGAGGTGACGAGGGGATGTCCGAAATCA 1371
Db      |||
QY 485 LeuAlaAspLeuGlyThrSerGlnHisGluLeuGluArgGlyAspValProLysAla 504
Db      |||
QY 1372 CTTGAGTGTACATGAGTACTACATGCTCGGAGCGGCGGAGCGGCGGAGCAGTGA 1431
Db      |||
QY 505 IleGlnCysTyrMetAsnAspThrAsnAlaSerGluArgGluAlaValGluHisValLys 524
Db      |||
QY 1432 TGGCTGATAGCGGAGGTGTGGAAGAGATGAATGCGGAGAGGTGTCGAGGATTCCTCCA 1491
Db      |||
QY 525 PheLeuIleArgGluAlaTrpLysGluMetAsnThrValThrAlaSerAspCysPro 544
Db      |||
QY 1492 TTCGGCAAGATTTTATGAGTGTGCGAGTTCGATTTAGGAAGATGCGCGCAGTTGATGTAC 1551
Db      |||
QY 545 PheThrAspLeuValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 564
Db      |||
QY 1552 CATATGAGAGTGGCAGCGCACACACCCCTATTATACATCAACAATGACCAAGAAC 1611
Db      |||
QY 565 LeuAspGlyAspGlyHisGlyValGlnHisSerGluIleHisGlnGlnMetGlyGlyLeu 584
Db      |||
QY 1612 TTATTTCGAGCCCTTT 1626
Db      |||
QY 585 LeuPheGlnProTyr 589

```

Search completed: February 25, 2004, 06:00:37
Job time : 64 secs

Alignment Scores:
Pred. No.: 1,36e-260 Length: 543
Score: 2881.00 Matches: 543
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.07% Indels: 0
DB: Gaps: 0

US-09-938-956-6 (1-1632) x US-09-938-956-7 (1-543)

QY 1 ATGAGACGATCCGGAATACACCTCTCGTTGGAGTCAATCTCATCAATCGCTT 60
DB 1 MetArgSerGlyAsnTyrAsnProSerArgPaspValAsnPhellegInSerLeu 20
QY 61 CTCAGTGACTATAGGAGGACAAACACGTCATTAGGCTTCAGCTGCTCACTTTGGTG 120
DB 21 LeuSerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuVal 40
QY 121 AGATCGAATCGAGAAAGAAACGATCAATTCGACAACTTGAGTGGATCGATCACTTG 180
DB 41 LysMetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspLeu 60
QY 181 CAGAGGATGGGCTGCTCCGATCAATTCCAAATAGCTTCAAGAAATCTTGTCTCTATA 240
DB 61 GlnArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerile 80
QY 241 TATCTCGACCATCACTATTACAGAACCTTTTCCAAAGAGAAAGGATCTCTACTCC 300
DB 81 TyrLeuAspHisTyrLysAsnProPheProLysGluGluArgAspLeuTyrSer 100
QY 301 ACATCTCTGATTTAGGCTCCCTCAGAGAACATGTTTCAAGTCGCAACAGAGTATTC 360
DB 101 ThrSerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPhe 120
QY 361 GATAGTTTCAGAACAGAGGGTGTGATTCAGAAAGCTTTAGCGACACACAGAGA 420
DB 121 AspSerPheLysAsnGluGluGlyPheLysGluSerLeuSerAspAspThrArgGly 140
QY 421 TTCTTGCACTGATCAAGCTTCCTTTCTGTTGACGGAAGGCAACACCGCTCCAGTCA 480
DB 141 LeuLeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSer 160
QY 481 GCGAGGGAATCCGCAACCAATTTTGGAGGAAAGTGAACGAGGGTGGTGTGATGCG 540
DB 161 AlaArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyValAspGly 180
QY 541 GACCTTTTACAGAAATCGCATATCTTTGGACATCCCTCTTATTTGGAGGATTAAGG 600
DB 181 AspLeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArg 200
QY 601 CCAATGCACTGTGTGATCGAATGTATAGGAGGCGCCGACATGAATCCAGTAGTG 660
DB 201 ProAsnAlaProValTrpIleGluTyrTrpArgLysArgProAspMetAsnProValVal 220
QY 661 TTGAGGTTCCGATCTGCTTAATATATTTTCAAGCAATTTCAAGAGGCTCAA 720
DB 221 LeuGluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluLeuLys 240
QY 721 GAATCTCTCAGTGTGGAGAAATCTGCGTTTGTGAGAGCTGCGCTTCGCAAGGAT 780
DB 241 GluSerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAsp 260
QY 781 AGACTGGTGAATGCTACTTTTGGAAATATCTGGATCATCGGACCGATCGATCGAAGT 840
DB 261 ArgLeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSer 280
QY 841 GCAGGATATGATGGCAAGTCAAGCTCTGATTACGTCATGATCATATATATGAT 900
DB 281 AlaArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspIleTyrAsp 300
QY 901 GTCTATGGCACTTAGAAGAACTCGAACATTCACCTGATTCATTCGAGATGGGATATA 960

DB 301 ValTyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgArgTrpAspIle 320
QY 961 AACTCAATCGACCAACTTCCGATTTACATCAACTGTGCTTTTTCGCACTCAACAATTC 1020
DB 321 AsnSerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAspPhe 340
QY 1021 GTCGATCATACATCGTACGATGTTATGAAGAGAAAGCGCTCAAGCTTATACCTCCTG 1080
DB 341 ValAspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeu 360
QY 1081 CCGCAATCGTGGTGTGATTTGGCGGATAGTATATGTTAGAGCAGCGTGTCTACGGC 1140
DB 361 ArgGlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGly 380
QY 1141 GCGCACAAACCAAGTTTGAAGAGTATTTGAGAACTATCGGACAGTCAAGTGGGCCCC 1200
DB 381 GlyHisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyPro 400
QY 1201 TGATGTAAACGACATATTTCCGAGTAACAGATTCGTTCCAAAGAGAGACCGTCGAC 1260
DB 401 CysMetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAsp 420
QY 1261 AGTTTGTACAAATACACGATTTAGTTCGTTGTCATCTCTGTTCTGCGGCTTCTGAT 1320
DB 421 SerLeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAsp 440
QY 1321 GATTTCGGAACCTCGTGGAGAGCTGAGCAGAGGGGATGTCGCGAAATCACTTCAGTGC 1380
DB 441 AspLeuGlyThrSerValGluValSerArgGlyAspValProLysSerLeuGlnCys 460
QY 1381 TACATGAGTCACTACAAATCATCGAGCGGAGCGCGAGCAGCAGTGAATGGCTGATA 1440
DB 461 TyrMetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIle 480
QY 1441 GCGAGGTGTGGAGAGAGATGAATGCGAGAGGGTGTGAGAGGATTCCTCATTCGCAAA 1500
DB 481 AlaGluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLys 500
QY 1501 GATTTTATAGATGTGAGTTGATTTAGGAGGATGCGCAGTGTGATGTACCATATGGA 1560
DB 501 AspPheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGly 520
QY 1561 GATGGGACGCGCACACACCTTATATATATATATATATATATATATATATATATATAT 1620
DB 521 AspGlyHisGlyThrGlnHisProIleIleHisGlnMetThrArgThrLeuPheGlu 540
QY 1621 CCCTTTGCA 1629
DB 541 ProPheAla 543

RESULT 2

US-09-887-586A-22
; Sequence 22, Application US/09887586A
; Patent No. US20020094556A1
; GENERAL INFORMATION:
; APPLICANT: Chapell, Joseph P.
; APPLICANT: No. US20020094556A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 599

TYPE: PRT
ORGANISM: Mentha spicata
US-09-887-586A-22

Alignment Scores:
Pred. No.: 4,17e-260 Length: 599
Score: 2876.00 Matches: 542
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.90% Indels: 0
DB: 9 Gaps: 0

US-09-938-956-6 (1-1632) x US-09-887-586A-22 (1-599)

QY 4 AGAGCATCCGAACTACAACTCTCTGTTGGATGTCACCTTCATCCCATCGCTCTC 63
Db 58 ArgArgSerGlyAsnProSerArgTrpAspValAsnPheIleGlnSerLeuLeu 77
QY 64 AGTGACTATAAGGAGGACAAACAGCTGATTAGGCGCTCTGAGCTGGTGCATCTTGGTGAAG 123
Db 78 SerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuValLys 97
QY 124 ATGGAACCTGGAGAAAGAAACGATCAATTCGACAACTTGAGTTGATCGATTCGAC 183
Db 98 MetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspAspLeuGln 117
QY 184 AGGATGGGGCTGCTCGATCATTTCCAAATCAGTTCAAAGAAATCTTGTCTCTATATAT 243
Db 118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluIleLeuSerSerIleTyr 137
QY 244 CTCGACCATCACTATTACAGAACCTTTTCCAAAGAAAGAAAGGATCTCTACTCCACA 303
Db 138 LeuAspHisIleTyrLysAsnProPheProLysGluGluAspLeuTyrSerThr 157
QY 304 TCTCTTGATTTAGCTCTCTCAGACATCGTTTCAAGTCGCACAGAGGATTTGAT 363
Db 158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177
QY 364 AGTTTCAAGAACGAGGAGGTGAGTTCAAAGAAAGCTTTAGCGACGACACAGAGGATTG 423
Db 178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197
QY 424 TTGCACTGTATGAGCTCTCTCTGTTGAGGAGCGGAAACACCGCTCGATCGAG 483
Db 198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217
QY 484 AGGGAATTCGCCACCAATTTTGGAGAAAGTGAACGAGGTGGTGTGATGGGAC 543
Db 218 ArgGluPheAlaThrLysPheLeuGluLysValAsnGluGlyValAspGlyAsp 237
QY 544 CTTTTTAACAGAAATCGCATATTTTGGACATCCCTCTTCATTTGAGGANTTAAAGGCCA 603
Db 238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro 257
QY 604 ANTGCACCTGTGATCGATGATAGGAGGCGCGCATGAATCCAGTATGTTG 663
Db 258 AsnAlaProValTrpIleGluTrpArgLysArgProAspMetAsnProValValLeu 277
QY 664 GAGCTTGCCATCTCGATCAATTAATTTCTCAAGCACAAATTTCAAGAGAGCTCAAGAA 723
Db 278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluLeuLysGlu 297
QY 724 TCCTTCAGTGTGGAGAAATCTGGTTTGTGAGAGCTGCCCTTCGCAAGGATAGA 783
Db 298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317
QY 784 CTGTTGGATGTCTCTTTTGGAAATCTCGGATCATCGAGCCACGCTCAGCATCGAGTCCA 843
Db 318 LeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla 337
QY 844 AGGATAATGATGGCAAGTCAACGCTCTGATACGGTGATCGATGATTTATGATGTC 903
Db 338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357

QY 904 TATGGACCTTAGAAGAACTCGAACAAATTCAGTACCTCATTCGAGATGGATATAAC 963
Db 358 TyrGlyThrLeuGluGluLeuGlnPheThrAspLeuIleArgArgTrpAspIleAsn 377
QY 964 TCAATCGACCAACTTCCCGATTACATGCAACTGTCTTTTTCGCACTCAACAACTTCGTC 1023
Db 378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnPheVal 397
QY 1024 GATGATACATCGTACGATGTTATGAGGAGAAAGCGGTACAGTTTATACCTTACCTCGG 1083
Db 398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417
QY 1084 CAATCGTGGTGTGATTTCGGCGATTAAGTATATAGGAGGACCGGTGTTCTACGCGCGG 1143
Db 418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437
QY 1144 CACAAACCAAGTTTGGAGAGTATTGGAAGAACTCATGGCAGTCTGATAGTGGGCGCTGT 1203
Db 438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457
QY 1204 ATGTTAACGCACATATTCTTCGAGTACAGATTTCGTTCCACAAAGGAGACCGTCCGACAGT 1263
Db 458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477
QY 1264 TTGTACAAATACACAGATTTCGTTGGTTCATCTCTGCTTCGCTCGGCTTCTGATGAT 1323
Db 478 LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497
QY 1324 TTGGGAACCTCGTGGAGAGCTGAGCAGAGGGATGTCGCAAAATCACTTCAGTGTCTAC 1383
Db 498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517
QY 1384 ATGAGTGACTCAATGTCATCGAGGCGGAGGCGGAGCACGTCGTAATGGTGTGATAGCG 1443
Db 518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla 537
QY 1444 GAGGTGTGAAGAGATGAATCGGAGAGGTGTGCAAGGATTCCTCATTCGCAAGAT 1503
Db 538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557
QY 1504 TTTATAGGATGTGCACTTCATTTAGGAAGATGGCGGAGTTCCTCATTAATGGAGAT 1563
Db 558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577
QY 1564 GGCGACGGCACACCAACCCCTATTATACATCAACAAATGACCAAGAACCTTATTCGAGCCC 1623
Db 578 GlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGluPro 597
QY 1624 TTTGCA 1629
Db 598 PheAla 599

RESULT 3

US-09-903-012-22
; Sequence 22, Application US/09903012
; Patent No. US20020094557A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094557A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262

```

; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Mentha spicata
US-09-903-012-22

Alignment Scores:
Pred. No.: 4,17e-260 Length: 599
Score: 2876.00 Matches: 542
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.90% Indels: 0
DB: 9 Gaps: 0

US-09-938-956-6 (1-1632) x US-09-903-012-22 (1-599)
QY 4 AGACGATCCGGAACACTACAAACCTTCTCGTTGGATGTCACATTCATCAATCGCTTCTC 63
Db 58 ArgArgSerGlyAsnTyrAsnProSerArgTrpAspValAsnPheileGlnSerLeuLeu 77
QY 64 AGTGACTATAAGAGAGACAAACACGATGATTAGCGCTTCTGAGCTGCTCACTTTGGTGAAG 123
Db 78 SerAspTyrLysGluAspLysHisValileArgAlaSerGluLeuValThrLeuValLys 97
QY 124 ATGGAACTGAGAAAGAAACCGATCAATTCGACAACTTGATCGATCGATCGATGAG 183
Db 98 MetGluLeuGluLysGluThrAspGlnileArgGlnLeuGluLeuileAspAspLeuGln 117
QY 184 AGGATGGGCTGCGATCATTTCCAAATGATGTTCAAGAAATCTTGCTCTATATAT 243
Db 118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluLeuSerSerileTyr 137
QY 244 CTGACCATCATATTACAAAGACCTTTTCCAAAGAAAGAGGATCTCTACTCCACA 303
Db 138 LeuAspHisHisTyrTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157
QY 304 TCTCTTGCTATTAGGCTCTCAGAGAACATGTTTCAAGTCGCGACAGAGGATTCGAT 363
Db 158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177
QY 364 AGTTTCAAGAACGAGGAGGTGAGTTCAAAGAAAGCCTTAGCGACGACACAGAGATTG 423
Db 178 SerPheLysAsnGluGluGlyPheLysGluSerLeuSerAspAspThrArgGlyLeu 197
QY 424 TTGCAACTGATGAAGCTCTCTTCTGTTGACGGAAGGCGAAACCGCTCGAGTCAGCG 483
Db 198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217
QY 484 AGGGAATTCGCCACCAAAATTTTGGAGGAAAGTGAACGAGGCTGCTGTTGATGCGAC 543
Db 218 ArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyValAspGlyAsp 237
QY 544 CTTTAAACAGATCGCATATCTTTGGACATCCCTCTTCATGAGGATTAAGGCCA 603
Db 238 LeuLeuThrArgileAlaTyrSerLeuAspPheProLeuHisTrpArgileLysArgPro 257
QY 604 AATGCACTGTGGATCGAATGTTAGGAAGGCGGACATGAATCAATCAATGATGTTG 663
Db 258 AsnAlaProValTrpIleGluTyrArgLysArgProAspMetAsnProValValLeu 277
QY 664 GAGCTTCCCATCTCGACTTAATATTTGTCAGACGCAATTCGAGAGAGCTCAAGAA 723
Db 278 GluLeuAlaileLeuAspLeuAsnileValGlnAlaGlnPheGlnGluLeuLysGlu 297
QY 724 TCCTTCAGGTGGGAGAAATCTGCGTGTGTTGAGAGCTGCGCTTCGCAAGGATAGA 783
Db 298 SerPheArgTrpTrpArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317
QY 784 CTGCTGGAATGCTACTTTTGGAACTACTGGGATCATCGAGCCAGCTGAGCATGCAATGCA 843

```

```

Db 318 LeuValGluCysTyrPheTrpAsnThrGlylleGluProArgGlnHisAlaSerAla 337
QY 844 AGGATAATGATGGGCAAGTCACACGCTCTGATTTACGCTGATCGATGATATTTATGATGTC 903
Db 338 ArgIleMetMetGlyLysValAsnAlaLeuileThrValileAspAspileTyrAspVal 357
QY 904 TATGGCACCTTAGAAGAACTCGAACAAATTCACCTGACCTCATTCGAGAGATGGGATATAAAC 963
Db 358 TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuileArgArgTrpAspIleAsn 377
QY 964 TCAATCGACCAACTTCCCGATTTACATGCAACTGTGCTTTCTTGCACTCAACAACTTCGTC 1023
Db 378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397
QY 1024 GATGATACATCGTACGATGTTTGAAGAGAGAAAGCGCTCAACGTTATACCTTACCTCGCG 1083
Db 398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuAsp 417
QY 1084 CAATCGTGGGTTGATTTGGCGGATAAGTATATGCTAGAGACGCGTGGTTCTACCGCGGG 1143
Db 418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437
QY 1144 CACAAACCAAGTTTGAAGAGATATTGGAGAACTCATGCGCAGTCGATAGTGGCCCTGT 1203
Db 438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457
QY 1204 ATGTTAACGACACATATTCTCCGAGTAACAGATTCTGTTCAAAAGAGAGACCGTCGACAGT 1263
Db 458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477
QY 1264 TTGTACAAATPACCAACGATTTAGTTGGTTCATCTCTGCTGCTGCGCTTGTCTGATGAT 1323
Db 478 LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497
QY 1324 TTGGACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTCGCGAAATCACCTTCAGTGTCTAC 1383
Db 498 LeuGlyThrSerValGluGluValSerArgLysValPheValProLysSerLeuGlnCysTyr 517
QY 1384 ATGAGTCACTACATGATCGAGGCGGAGCGCGGAGGAGCAGTGAATCGCTGATAGCG 1443
Db 518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuileAla 537
QY 1444 GAGTGTGGAAGAGATGATGCGGAGGGGTGTCGAGGATTCCTCATTTCGCAAAAGAT 1503
Db 538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557
QY 1504 TTTATAGATGTGCGATTTGATTTAGGAAGGATGCGCAGTGTGATGTACCAATATGAGAT 1563
Db 558 PheileGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577
QY 1564 GGGCAGCGCACACAAACCTTATATATATATATATATATATATATATATATATATATATAT 1623
Db 578 GlyHisGlyThrGlnHisProIleileHisGlnMetThrArgThrLeuPheGluPro 597
QY 1624 TTTGCA 1629
Db 598 PheAla 599

RESULT 4
US-09-900-797-22
; Sequence 22, Application US/09900797
; Publication No. US20030087406A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20030087406A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/900,797
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/398,395
; PRIOR FILING DATE: 1999-09-17

```

```
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Mentha spicata
US-09-900-797-22

Alignment Scores:
Pred. No.: 4,17e-260 Length: 599
Score: 2876.00 Matches: 542
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.90% Indels: 0
DB: 10 Gaps: 0

US-09-938-956-6 (1-1632) x US-09-900-797-22 (1-599)

QY 4 AGAGATCGGAAACTCAACCCCTTCGTGGGATGTCATCAATTCATCCATCGCTTCTC 63
Db 58 ArgArgSerGlyAsnTyrAsnProSerArgTrpAspValAsnPheIleGlnSerLeuLeu 77
QY 64 AGTGACTATAGGAGGACAAACAGCTGATTAGGCTTCTCAGCTGGTCACTTTGGTGAAG 123
Db 78 SerAspTyrLysGluAspLysHisValIleArgAlaSerGluLeuValThrLeuValLys 97
QY 124 ATGGAATCGGAGAAAGAAACGGATCAAAATTCGAACACTTCAGTTGATCGATGCTTCGAG 183
Db 98 MetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuIleAspAspLeuGln 117
QY 184 AGGATGGGCTCGCATCAATTCCTCAAAATCAGTTCAAGAAATCTGTCTCTATATAT 243
Db 118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluLeuSerSerIleTyr 137
QY 244 CTCGACCATCACTATTCAAGAACCTTTCCAAAGAAAGAGGATCTCTACTCCACA 303
Db 138 LeuAspHisHisTyrTyrLysAsnProPheProLysGluLysGluArgAspLeuTyrSerThr 157
QY 304 TCTCTTCGATTAAGCTTCCTCAGAGAACATGGTTTCAAGTCGACAGAGGATTCGAT 363
Db 158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177
QY 364 AGTTTCAGACGAGGAGGTGAGTTCAAGAAAGCCTTAGCGACGACACACAGAGGATTCG 423
Db 178 SerPheLysAsnGluGluGlyGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197
QY 424 TTGCAACTGTATGAGCTTCCTTTCTGTGACGAGGCGAAACCCAGCTCGAGTCAGCG 483
Db 198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGlyGluThrThrLeuGluSerAla 217
QY 484 AGGGAATTCGCCACCAATTTTCGAGGAAAGTGAACGAGGTGGTGTGATGGCGAC 543
Db 218 ArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyGlyValAspGlyAsp 237
QY 544 CTTTAAACAAGATCGCATATCTTTTGACATCCCTCTTCATTCGAGGATTAAGGCCCA 603
Db 238 LeuLeuThrArgIleAlaTyrSerLeuAspIleProLeuHisTrpArgIleLysArgPro 257
QY 604 AATGACCTGTGTGATCGAATGTATAGGAGAGGCCCGACATCAATTCAGTAGTGTTC 663
Db 258 AsnAlaProValTrpIleGluTrpTyrArgLysArgProAspMetAsnProValValLeu 277
QY 664 GAGCTTGCACTACGCTAAATATTGTTCAAGCAACAATTCAGAAAGAGCTCAAGAA 723
Db 278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluGluLeuLysGlu 297
QY 724 TCCTTCAGGTGGAGAAATPACTGGGTTTCTGAGAGCTGCCCTTCGCAAGGATAGA 783
Db 298 SerPheArgTrpTyrArgAsnThrGlyPheValGluLysLeuProPheAlaArgAspArg 317
```

```
784 CTGGTGGAAATGCTACTTTTGGAAATCTGGGATCATCGAGCCACGTCACATGCAAGTCA 843
Db 318 LeuValGluCysTyrPheTrpAsnThrGlyIleIleGluProArgGlnHisAlaSerAla 337
QY 844 AGGATAATGATGGGCAAAAGTCAACGCTCTGATTACCGTGCATCGATGATATTTATGATGC 903
Db 338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357
QY 904 TATGGCACTTAGAAGAACTCGAACAATTCATGACCTCATTCGAGATGGGATATAAAC 963
Db 358 TyrGlyThrLeuGluGluLeuGlnPheThrAspLeuIleArgArgTrpAspIleAsn 377
QY 964 TCAATCGACCAACTCCCGATTACATGCAACTGTCTTCTTCTGCACTCAAACTTCGTC 1023
Db 378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397
QY 1024 GATGATACATCGTACGATGTTATGAAGAGAAAGCGTCAACGTTATACCTTACCTGGG 1083
Db 398 AspAspThrSerTyrAspValMetLysGluLysValAsnValIleProTyrLeuArg 417
QY 1084 CAATCGTGGTGTGATTTGGCGGATAAGTATATGTTAGAGGCGACGCTGTTCTACGGCGG 1143
Db 418 GlnSerTrpValAspLeuAlaAspLysTyrMetValGluAlaArgTrpPheTyrGlyGly 437
QY 1144 CACAAACCAAGTTTGGAAAGATTTTGGAGAACTCATGCGCAGTCGATAAGTGGGCCCTGT 1203
Db 438 HisLysProSerLeuGluLutyrLeuGluAsnSerTrpGlnSerIleSerGlyProCys 457
QY 1204 ATGTTAAGCACAATATCTTCGAGTAACAGATTCGTTCAAAAGGAGACCGTCACAGT 1263
Db 458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477
QY 1264 TTGTAACAATACCCAGATTTAGTTGCTGTCATCTTCGTTCTCGCGCTTCTGCTGATGAT 1323
Db 478 LeuTyrLysTyrHisAspLeuValArgTrpSerSerPheValLeuArgLeuAlaAspAsp 497
QY 1324 TTGGAACTCGTGGTGGAAAGATGAGCAGAGGGGATGCGCGAATCATCTTCAGTGTCTAC 1383
Db 498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517
QY 1384 ATGAGTGTCAATGTCATCGAGCGCGAGCGCGGAGGAGCGTGAATGGCTGATAGCG 1443
Db 518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTrpLeuIleAla 537
QY 1444 GAGGTGTGGAAGAGATGAATGCGGAGAGGGTGTGCAAGGATTTCCATTGCGCAAGAT 1503
Db 538 GluValTrpLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557
QY 1504 TTTATAGGATGTCAGTTGATTTAGAGGATGCGCGAGTTCATGATGATGATGATGATGAT 1563
Db 558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577
QY 1564 GGGCAGCGGCACACAACCCCTATTATATCATCAACAATGACCAAGATTTATTCAGAGGCC 1623
Db 578 GlyHisGlyThrGlnHisProIleIleHisGlnMetThrArgThrLeuPheGluPro 597
QY 1624 TTTGCA 1629
Db 598 PheAla 599

RESULT 5
US-09-887-586A-24
; Sequence 24, Application US/09887586A
; Patent No. US20020094556A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094556A11, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
```

[REDACTED]

```

; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Salvia officinalis
; US-09-903-012-24

Alignment Scores:
Pred. No.: 9.03e-144      Length: 591
Score: 1633.50           Matches: 305
Percent Similarity: 75.96% Conservatives: 109
Best Local Similarity: 55.96% Mismatches: 116
Query Match: 56.17%      Indels: 15
DB: 9                     Gaps: 7

US-09-938-956-6 (1-1632) x US-09-903-012-24 (1-591)
QY 4 AGAGATCCGGAACACTACACCTCTCGTGGATGTCACCTTCATCCATCGCTTCTC 63
DB 57 ArgargThrGlyGlyGlnProThrLeuTrpAspPheSerThrIleGlnLeuPheAsp 76
QY 64 AGTCACTATAAGGAGGACAAACACGCTGATAGGCTTCTGAGCTTCTGAGTGGTGAAG 123
DB 77 SerGluTrpLysGluGluLysHisLeuMetArgAlaAlaGlyMetIleAlaGlnValAsn 96
QY 124 ATGGAATCGGAGAGAAACGATCAATTCGACAACTTCGAGTTCGATGATCGATCGAG 183
DB 97 MetLeuLeuGlnGluValAspSerIleGlnArgLeuGluLeuLeuLeuAspLeuArg 116
QY 184 AGGATGGGCTGTCGATCATTTCCAAATAGATTCAGAAATCTTGCTCTATATAT 243
DB 117 ArgLeuGlyIleSerCysHisPheAspArgGluLeuValGluLeuAsnSerLysTyr 136
QY 244 CTCGACCATCACTATTACAAGAACCTTTTCCAAAGAAAGAGGATCTCTACTCCACA 303
DB 137 -----TyrThrAsnAsnGluLeu-----AspGluSerAspLeuTyrSerThr 150
QY 304 TCTCTTCGATTTAGCTCTCAGAGAACATGGTTTTCAGTCCGACAGAGGATTCGAT 363
DB 151 AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp 170
QY 364 AGTTTCAAGAACGAGGAGGGT---GAGTTCAGAAAGAGCTTTAGCGACGACACACAGGA 420
DB 171 CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190
QY 421 TTGTTGCACTGTATGAGCTCTCTTCTGTTGACGGAAGGCGAAACACGCTCGAGTCA 480
DB 191 LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu 210
QY 481 CGGAGGGAATTCGCCCAATTTTTCGAGGAGAAAGTGAACGAGGGTGTGTTGATGGC 540
DB 211 AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu-----ValAspLys 227
QY 541 GAC-----CTTTTAAAGAAATCGCATTTCTTTGGACATCCCTCTTCATTTGAGGATT 594
DB 228 AspIleAsnLeuLeuSerSerIleGluArgAlaLeuLeuGluLeuProThrHisTrpArgVal 247
QY 595 AAAGGCCAAATGCACTGTGATGATCAATGGTATAGGAGAGGCCCGCATGAAATCCA 654
DB 248 GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267

```

```

QY 655 GTAGTGTGGAGCTTGCATACCTCGACTTAATATTGTTCAAGCACAAATTTCAAGAGAG 714
DB 268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlnGlu 287
QY 715 CTCAAAGAATCCTCAGGTGGTGGAGAAATCTAGTGGTGGTGGTGGTGGTGGTGGTGG 774
DB 288 LeuLysGluAlaSerArgTrpTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307
QY 775 AGGATAGACTGGTGAATCTACTTTTGGAAATCTGGATCATCGATCATCGACCGTCCAGAT 834
DB 308 ArgAspArgIleValGluCysTyrTrpTrpThrThrGlyValValGluArgGluHis 327
QY 835 GCAAGTGCAGGATTAATGATGGCAAGTCAACGCTCTGATTAACGCTGATCATGATATT 894
DB 328 GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrThrIleAspVal 347
QY 895 TATGATGCTATGGCACCTTAGAAGAACTCGAACAATTCACCTGACCTCATTCGAGAGTGG 954
DB 348 PheAspIleTyrGlyThrLeuGluLeuLeuPheThrThrAlaIleGlnArgTrp 367
QY 955 GATATAAACTCAATCGACCAACTTCCGATTCACGAACTGTGCTTCTTCTGCACTCAAC 1014
DB 368 AspIleGluSerMetLysGlnLeuProProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387
QY 1015 AACTTCGTGATGATACATCGTACGATGTTTGAAGAGTATTGGAGAACTCATGCGAGTCA 1074
DB 388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407
QY 1075 TACCTCGGCACTGCTGGTGTGATTTGGCGGATAGTATATGTTAGAGCAGCGTGGTTC 1134
DB 408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTrpTyr 427
QY 1135 TACGCGGCGCACAAACCAAGTTTGAAGAGTATTGGAGAACTCATGCGAGTCAAGTATG 1194
DB 428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447
QY 1195 GGGCCCTGTATGTTAACGCACATATTTCCGAGTAACAGATTTCGTTCAACAGAGAGACC 1254
DB 448 GlyIleProIleLeuSerHisLeuPheArgLeuThrAspSerIleGluGluAsp 467
QY 1255 GTCCGACAGTTTTCACAAATACCACTTTCGTTGCTTTCGTTTCATCTTCTGCGGCTT 1314
DB 468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487
QY 1315 GCTGATGATTTGGGAACCTCGTGGAGAGGTGACGAGGAGGATGCGCGAATCACTT 1374
DB 488 AlaAspAspMetGlyThrSerLeuAspGluValGluArgLysValProLysSerVal 507
QY 1375 CAGTGTCTACATGAGTACATACCAATTCACGAGGCGGAGCGCGAAGCACGCTGAATGG 1434
DB 508 GlnCysTyrMetAsnGluLysAsnAlaSerGluGluGluAlaArgGluHisValArgSer 527
QY 1435 CTGATAGCGGAGGTGTGGAAGAGATGATGATGCGGAGAGGAGGTGCGAAGATTCTCCATTC 1494
DB 528 LeuIleAspGlnThrTriplysMetMetAsnLysGluMetMetThr-----SerSerPhe 545
QY 1495 GCGAAGATTTTATAGGATGTCAGTGTGATTTAGGAAGGATGCGCGAGTGTGATGAC--- 1551
DB 546 SerLysTyrPheValGlnValSerAlaLeuLeuAlaArgMetAlaGlnTrpIleTyrGln 565
QY 1552 CATATAGGAGATGGGCGACCGCACACACACCTTATATACATCAACAATAGCCAGAAC 1611
DB 566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585
QY 1612 TTATTCGAGCCCTTT 1626
DB 586 LeuPheAspArgTyr 590

```

RESULT 7
 US-09-900-797-24
 ; Sequence 24, Application US/09900797
 ; Publication No. US20030087406A1


```

GENERAL INFORMATION:
APPLICANT: Chappell, Joseph
APPLICANT: No. US20030087406A11, Joseph P.
APPLICANT: Stark, Courtney M.
APPLICANT: Manna, Kathleen R.
TITLE OF INVENTION: SYNTHASES
FILE REFERENCE: 07678-025001
CURRENT APPLICATION NUMBER: US/09/900,797
CURRENT FILING DATE: 2001-07-06
PRIOR APPLICATION NUMBER: US/09/398,395
PRIOR FILING DATE: 1999-09-17
PRIOR APPLICATION NUMBER: 60/130,628
PRIOR FILING DATE: 1999-04-22
PRIOR APPLICATION NUMBER: 60/150,262
PRIOR FILING DATE: 1999-08-23
NUMBER OF SEQ ID NOS: 58
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 24
LENGTH: 591
TYPE: PRT
ORGANISM: Salvia officinalis
US-09-900-797-24

Alignment Scores:
Pred. No.: 9,036-144 Length: 591
Score: 1633.50 Matches: 305
Percent Similarity: 75.96% Conservative: 109
Best Local Similarity: 55.96% Mismatches: 116
Query Match: 56.17% Indels: 15
DB: 10 Gaps: 7

US-09-938-956-6 (1-1632) x US-09-900-797-24 (1-591)
QY 4 AGAGATCCGGAACTACAAACCTTCCTCGTGGAGTGCATTCATCAATCGCTTCTC 63
Db 57 ArgArgThrGlyGlyTyrGlnProThrLeuTrpAspPheSerThrIleGlnLeuPheAsp 76
QY 64 AGTGACTATAAGGAGACAAACACGATGATTAGGGCTTCGAGCTCGCTCACTTTGGTGAAG 123
Db 77 SerGluTyrIysGluGlnLysHisLeuMetArgAlaAlaGlyMetIleAlaGlnValAsn 96
QY 124 ATGGAACTGGAGAAAGAACGGATCAAAATTCGACAACTTGAGTGCATCGATCTGCAG 183
Db 97 MetLeuLeuGlnGluGluValAspSerIleGlnArgLeuGluLeuIleAspAspLeuArg 116
QY 184 AGGATGGGCTGTCATCATTTCCAAATGAGTTCAAGAAATCTTGCTCTATATAT 243
Db 117 ArgLeuGlyIleSerCysHisPheAspArgGluIleValGluIleLeuAsnSerLysTyr 136
QY 244 CTCGACCATCACATATTACAAAGACCTTTTCCAAAGAAAGAGGATCTCTACTCCACA 303
Db 137 -----TyrThrAsnAsnGluIle-----AspGluSerAspLeuTyrSerThr 150
QY 304 TCCTCTGCATTTAGGCTCTCAGAAACATGGTTTCAAGTGCACAGAGTATTCGAT 363
Db 151 AlaLeuArgPheLysLeuLeuArgGlnTyrAspPheSerValSerGlnGluValPheAsp 170
QY 364 AGTTTCAAGAACAGAGAGGCT--GAGTTCAAGAAAGCTTAGCGACACACACAGAGA 420
Db 171 CysPheLysAsnAspLysGlyThrAspPheLysProSerLeuValAspAspThrArgGly 190
QY 421 TTGTTGCACTGTATGAGCTTCCTTTCTGTTGACGGAAGGGGAAACCCACGCTCGAGTCA 480
Db 191 LeuLeuGlnLeuTyrGluAlaSerPheLeuSerAlaGlnGlyGluGluThrLeuHisLeu 210
QY 481 GCGAGGGAATCCCAACCAATTTTGGAGGAAAGAGTGAACGAGGCTGGTGTGATGCC 540
Db 211 AlaArgAspPheAlaThrLysPheLeuHisLysArgValLeu-----ValAspLys 227
QY 541 GAC-----CTTTTAAACAAGATCGCATATTTCTTGACATCCCTCTTCATGGAGGATT 594
Db 228 AspIleAsnLeuLeuSerSerIleGluArgAlaLeuGluLeuProThrHisTrpArgVal 247

```

RESULT 8

```

QY 595 AAAAGGCCAAATCGACCTGTGTGGATCGAATGTTATAGGAAGAGGCCCGACATGAATCCA 654
Db 248 GlnMetProAsnAlaArgSerPheIleAspAlaTyrLysArgArgProAspMetAsnPro 267
QY 655 GTAGTGTGGAGCTTGGCACTACTCGCTAAATATTGTTCAAGCACAATTTCAAGAAGAG 714
Db 268 ThrValLeuGluLeuAlaLysLeuAspPheAsnMetValGlnAlaGlnPheGlnGlnGlu 287
QY 715 CTCAAAGAATCCTTTCAGGTGGTGGAGAAATACATCGGCTTTGTTGAGAAGCTGCCCTTCGCA 774
Db 288 LeuLysGluAlaSerArgTrpTrpAsnSerThrGlyLeuValHisGluLeuProPheVal 307
QY 775 AGGATAGACTGTGTGAATCTACTTTTGGAAATCTGGAATCTGGAATCTGAGCCACCTCAGCAT 834
Db 308 ArgAspArgIleValGluCysTyrTyrTrpThrThrGlyValValGluArgArgGluHis 327
QY 835 GCAAGTGCAGAGATAATGATGGGCAAGTCAACGCTCTGATTACGGTGTGATGATGATATT 894
Db 328 GlyTyrGluArgIleMetLeuThrLysIleAsnAlaLeuValThrThrIleAspAspVal 347
QY 895 TATGATGCTTATGGCACCTTAGAAGAACTGCAACAATTCATGACCTCATTCGAAGATGG 954
Db 348 PheAspIleTyrGlyThrLeuGluLeuGluLeuPheThrThrAlaIleGlnArgTrp 367
QY 955 GATATAAATCAATCGACCAACTTCCGATTACATGCACTGCTGCTTTCTTGCACTCAAC 1014
Db 368 AspIleGluSerMetLysGlnLeuProProTyrMetGlnIleCysTyrLeuAlaLeuPhe 387
QY 1015 AACTTCGTGATGATCATCATCGTACGATGTTATGAAGAGAAAGCGGTCAACGTTATACCC 1074
Db 388 AsnPheValAsnGluMetAlaTyrAspThrLeuArgAspLysGlyPheAsnSerThrPro 407
QY 1075 TACTCGGGCAATCGTGGTGTGATTGGCGGATAAGTATATATGTTAGAGGACCGTGGTTC 1134
Db 408 TyrLeuArgLysAlaTrpValAspLeuValGluSerTyrLeuIleGluAlaLysTrpTyr 427
QY 1135 TACGGCGGACAAACCAACTTTGGAAGAGTATTTCGAGAACTCATGCGAGTGCATTAAGT 1194
Db 428 TyrMetGlyHisLysProSerLeuGluGluTyrMetLysAsnSerTrpIleSerIleGly 447
QY 1195 GGGCCCTGTATGTAAACGACATATCTTCCGATCAACGATTCGTTCAAAAGGAGACC 1254
Db 448 GlyIleProIleLeuSerHisLeuPheArgLeuThrAspSerIleGluGluGluAsp 467
QY 1255 GTCACAGATTGTACAATAACACGATTAGTTCGTTGGTTCATCCTTCGTTCTGCGGCTT 1314
Db 468 AlaGluSerMetHisLysTyrHisAspIleValArgAlaSerCysThrIleLeuArgLeu 487
QY 1315 GCTGATGATTGGAACTCGTGGAAAGAGTGAATCGGAGAGAGGTGTCGAAGGATTCCTCATTC 1374
Db 488 AlaAspAspMetGlyThrSerLeuAspGluValGluArgGlyAspValProLysSerVal 507
QY 1375 CAGTGTACATGAGTACTACATGATCGAGGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAG 1434
Db 508 GlnCysTyrMetAsnGluLysAsnAlaSerGluGluGluAlaArgGluHisValArgSer 527
QY 1435 CTGATAGCGGAGGTGTGGAAGAAAGATGAATCGGAGAGAGGTGTCGAAGGATTCCTCATTC 1494
Db 528 LeuIleAspGlnThrTrpLysMetAsnLysGluMetMetThr-----SerSerPhe 545
QY 1495 GCGAAAGATTTTATAGGATGTGCAGTGTATTAGGAAGGATGGCGCAGTTGATGTAC--- 1551
Db 546 SerLysTyrPheValGlnValSerAlaAsnLeuAlaArgMetAlaGlnTrpIleTyrGln 565
QY 1552 CATATGAGATGGGACCGCACACAACACCTTATTATCATCAACAAATGACCAAGACACC 1611
Db 566 HisGluSerAspGlyPheGlyMetGlnHisSerLeuValAsnLysMetLeuArgGlyLeu 585
QY 1612 TTATTCGAGCCCTTT 1626
Db 586 LeuPheAspArgTyr 590

```

US-09-987-586A-54
; Sequence 54, Application US/09887586A
; Patent No. US20020094556A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094556A11, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 54
; LENGTH: 590
; TYPE: PRT
; ORGANISM: Salvia officinalis
US-09-887-586A-54

Alignment Scores:
Pred. No.: 2,79e-136 Length: 590
Score: 1553.50 Matches: 291
Percent Similarity: 72.48% Conservative: 104
Best Local Similarity: 53.39% Mismatches: 141
Query Match: 53.42% Indels: 9
Gaps: 3

US-09-938-956-6 (1-1632) x US-09-887-586A-54 (1-590)

QY	1	ATGAGACATCGGAAATACACACCTTCTGTTGGGATGTCACATTCATCCATCGCTT	60
DB	51	IleArgSerGlyAspTyrGlnProSerLeuIrpAspPheAsnTyrIleGlnSerLeu	70
QY	61	CTCAGTACTATAAGGAGGACAAACACGTCGATTTAGGGCTTCGAGCTGCTCACCTTGGTG	120
DB	71	AsnThrProTyrLysGluGlnArgHisPheAsnArgGlnAlaGluLeuLeuMetGlnVal	90
QY	121	AGATGGAACTGGGAAAGAAACGGATCAATTCGACAACTTGATTCGATCGATGCTTG	180
DB	91	ArgMetLeuLysValLysMetGluAlaIleGlnGlnLeuGluLeuLeuLeuLeuLeuLeu	110
QY	181	CAGAGATGGGCTGTCGATCATTTCCAAATAGAGTTCAAGAAATCTTCTCTCTATA	240
DB	111	GlnTyrLeuGlyLeuSerTyrPhePheGlnAspGluIleLysGlnIleLeuSerSerIle	130
QY	241	TATCTCGACCATCACTATTACAGAACCCCTTTTCCAAAGAGAAAGAGGATCTCTACTCC	300
DB	131	HisAsnGluProArgTyrPheHisAsn-----AsnAspLeuTyrPhe	144
QY	301	ACATCTCTGCTTTAGGCTCTCAGACACATGTTTCAAGTCGCACAGAGTATTC	360
DB	145	ThrAlaLeuGlyPheArgIleLeuArgGlnHisGlyPheAsnValSerGluAspPhe	164
QY	361	GATAGTTTCAAGAACGAGGAG--GCTGAGTTTCAAGAAAGCCTTTAGCGACGACACACAGA	417
DB	165	AspCysPheLysIleGluLysCysSerAspPheAsnAlaAsnLeuAlaGlnAspThrLys	184
QY	418	GAATGTTGCACTGATGATGCTCTTCTGTTGACGGAAGGCGGAACACGCTCGAG	477
DB	185	GlyMetLeuGlnLeuTyrGluAlaSerPheLeuLeuArgGluGlyGluAspThrLeuGlu	204
QY	478	TCAGCGAGGGAATTCGCCACCAATTTTGGAGGAAAGTCAACGAGGCTGGT-----	531
DB	205	LeuAlaArgPheSerThrArgSerLeuArgGluLysPheAspGluGlyGlyAspGlu	224
QY	532	GTTGATGGCGACCTTTTAAACAAATCGCATATTTCTTTGGACATCCCTCTTCATTGGAGG	591

DB	225	IleAspGluAspLeuSerSerTrpIleArgHisSerLeuAspLeuProLeuHisTrpArg	244
QY	592	ATTAAAGGCCAAATGCACCTGTGTGGATCGAATGGTATAGGAAGGCGCCGACATGAT	651
DB	245	ValGlnGlyLeuGluAlaArgTrpPheLeuAspAlaTyrAlaArgProAspMetAsn	264
QY	652	CCAGTAGTGTTCGAGCTTCCCATCTCGACTTAAATATTGTTCAAGCAAAATTTCAAGAA	711
DB	265	ProLeuIlePheLysLeuAlaLysLeuAsnPheAsnIleValGlnAlaThrTyrGlnGlu	284
QY	712	GAGCTCAAGAATCCTTCAGGTGGTGGAGAAATACTGGGTTTGTGTGAGAGCTGCCCTTC	771
DB	285	GluLeuLysAspIleSerArgTrpTrpAsnSerCysLeuAlaGluLysLeuProPhe	304
QY	772	GCAAGGATAGACTGTGTGAATGCTACTTTTGGAAATCTGGGATCATCCAGGACCATCTAG	831
DB	305	ValArgAspArgIleValGluCysPhePheTrpAlaIleAlaPheGluProHisGln	324
QY	832	CATGCAAGTCAAGGATAATGATGGCAAAAGTCAACGCTCTGATTACGCTGATCGATGAT	891
DB	325	TyrSerTyrGlnArgLysMetAlaAlaValIleThrPheIleThrIleAspAsp	344
QY	892	ATTATGATGCTGTGGCACCTTAGGAACCTCGAACAATTCAGCTGACCTCATTTCCAAGA	951
DB	345	ValTyrAspValTyrGlyThrIleGluGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu	364
QY	952	TGGGATATAAATCAATCGACCACTTCCCGATTACATCAACTGTGCTTTCTTTCGACTC	1011
DB	365	TrpAspAsnLysSerIleSerGlnLeuProTyrTyrMetGlnValCysTyrLeuAlaLeu	384
QY	1012	AACAATTCGTGATGATACATCGTACGATGTATGATGAGGAGAAAGCGTCAACGTTATA	1071
DB	385	TyrAsnPheValSerGluArgAlaTyrAspIleLeuLysAspGlnHisPheAsnSerIle	404
QY	1072	CCCTACTCGGCAATCGTGGTGTGATTTGGCGGATAAGTATATGTTAGAGCGACCGGTGG	1131
DB	405	ProTyrLeuGlnArgSerTrpValSerLeuValGluGlyTyrLeuLysGluAlaTyrTrp	424
QY	1132	TTCTAGCGCGGCAAAACCAAGTTTGGAGAGTATTTCGAGAACTCATCGGAGTGCATA	1191
DB	425	TyrTyrAsnGlyTyrLysProSerLeuGluGluTyrLeuAsnAsnAlaLysIleSerIle	444
QY	1192	AGTGGGCGCTGTATGTTAAGCACATATCTTCCGAGTACACATTCGTTTCACAAAGGAG	1251
DB	445	SerAlaProThrIleLeuSerGlnLeuTyrPheThrLeuAlaAsnSerIleAspGluThr	464
QY	1252	ACCGTCGACATTTGTACAAATACCGATTTAGTTCGTTTGGTTCATCCTTCGTTCTCGG	1311
DB	465	AlaIleGluSerLeuTyrGlnTyrHisAsnIleLeuTyrLeuSerGlyThrIleLeuArg	484
QY	1312	CTTGTCTGATGTTGGGAACTCGGTGGAGAGGTGACGAGCGGATGTCGGAATCA	1371
DB	485	LeuAlaAspAspLeuGlyThrSerGlnHisGluLeuGluArgGlyAspValProLysAla	504
QY	1372	CTTCAGTGTACATGATGACTACAATTCGATCGAGCGGAGCGCGGAGACGCGTGA	1431
DB	505	IleGlnCysTyrMetAsnAspThrAsnAlaSerGluArgGluAlaValGlnHisValLys	524
QY	1432	TGGCTGATACGAGGTGTGGAGAGAGTCAATCGGAGGAGGTGTCGAAGGATTCCTCA	1491
DB	525	PheLeuIleArgGluAlaTrpLysGluMetAsnThrValThrThrAlaSerAspCysPro	544
QY	1492	TTCCGCAAGATTTTATGATGTGTGAGTGTGATTTAGGAGGATGCGCGCATGATGATC	1551
DB	545	PheThrAspAspLeuValAlaAlaAlaAsnLeuAlaArgAlaAlaGlnPheIleTyr	564
QY	1552	CATAATGAGATGGGACGCGCACAAACCCCTATTATATACATCAACAAATGACGAGACC	1611
DB	565	LeuAspGlyAspGlyHisGlyValGlnHisSerGluIleHisGlnGlnMetGlyLeu	584
QY	1612	TTATTTCGAGCCCTTT 1626	

Db 585 LeupheGlnProTyr 589

RESULT 9

US-09-903-012-54
 / Sequence 54, Application US/0903012
 / Patent No. US20020094557A1
 / GENERAL INFORMATION:
 / APPLICANT: Chappell, Joseph
 / APPLICANT: No. US20020094557A1, Joseph P.
 / APPLICANT: Starks, Courtney M.
 / APPLICANT: Manna, Kathleen R.
 / TITLE OF INVENTION: SYNTHASES
 / FILE REFERENCE: 07678-025001
 / CURRENT APPLICATION NUMBER: US/09/903.012
 / CURRENT FILING DATE: 2001-07-11
 / PRIOR APPLICATION NUMBER: 09/398,395
 / PRIOR FILING DATE: 1999-09-17
 / PRIOR APPLICATION NUMBER: 60/100,993
 / PRIOR FILING DATE: 1998-09-18
 / PRIOR APPLICATION NUMBER: 60/130,628
 / PRIOR FILING DATE: 1999-04-22
 / PRIOR APPLICATION NUMBER: 60/150,262
 / PRIOR FILING DATE: 1999-08-23
 / NUMBER OF SEQ ID NOS: 58
 / SOFTWARE: FastSeq for Windows Version 3.0
 / SEQ ID NO 54
 / LENGTH: 590
 / TYPE: PRT
 / ORGANISM: Galvia officinalis
 / US-09-903-012-54

Alignment Scores:

Pred. No.: 2,79e-136 Length: 590
 Score: 1553.50 Matches: 291
 Percent Similarity: 72.48% Conservative: 104
 Best Local Similarity: 53.39% Mismatches: 141
 Query Match: 53.42% Indels: 9
 Db: 9 Gaps: 3

US-09-938-956-6 (1-1632) x US-09-903-012-54 (1-590)

QY 1 ATGAGACGATCCGGAACCTCAACCTTCGTGGGATCTCAACTTCATCCATCGCTT 60
 Db 51 lIeArgSerGlyApyrGlnProSerLeuTrpAsnSerPheAsnTrpGlnSerLeu 70
 QY 61 CTCAGTACTATAAGGAGCAACACAGTCGATAGGCTTCTGAGCTGTCATCTTGGTG 120
 Db 71 AsnThrProTyrLysGluGlnArgHisPheAsnArgGlnAlaGluLeuMetGlnVal 90
 QY 121 AAGATGGAACGGAAGAAACGGATCAATTCGACACTTCAGTTCGATCGACTTG 180
 Db 91 ArgMetLeuLysValLysMetGluAlaLeuGlnLeuLysGlnLeuLeuLeuLeu 110
 QY 181 CAGAGGATGGGCTGTCGATCAATTCGAAATGAGTTCGAAATCTTGTCTCTATA 240
 Db 111 GlnTyrLeuGlyLeuSerTyrPhePheGlnAspGluLysGlnLeuLeuSerLeu 130
 QY 241 TATCTCGACCATCACTATTCAAGAACCTTTTCCAAAGAAAGAGGATCTCTACTCC 300
 Db 131 HisanGluProArgTyrPheHisasn-----AsnAspLeuTyrPhe 144
 QY 301 ACATCTCTGATTTAGGCTCTCAGAGAACATGGTTTCAATCGCACAGAGGTATTC 360
 Db 145 ThrAlaLeuGlyPheArgIleLeuArgGlnHisGlyPheAsnValSerGluAspValPhe 164
 QY 361 CATAGTTTCAAGAACGAGGAG--GGTGGTTCGAAAGAGCCTTAGCCGACCAACGAGA 417
 Db 165 AspCysPheIleGluLysCysSerAspPheAsnAlaAsnLeuAlaGlnAspThrLys 184
 QY 418 GGATTTGCAACTGTATGAGCTTCTTCTGTGACGGAAGCGGAAACCAACGCTCGAG 477
 Db 185 GlyMetLeuGlnLeuTyrGluAlaSerPheLeuLeuArgGluGlyLeuAspThrLeuGlu 204

QY 478 TCACCGAGGGAATTCGCCACCAAAATTTTGGAGGAAAAAGTAGTGAACGAGGCTGGT----- 531
 Db 205 LeuAlaArgArgPheSerThrArgSerLeuArgGluLysPheAspGluGlyGlyAspGlu 224
 QY 532 GTTCATGGCGACCTTTTAAACAAGATCGCATATTTCTTGGACATCTCTTTCATTTGGAGG 591
 Db 225 IleAspGluAspLeuSerSerTrpIleArgHisSerLeuAspLeuProLeuHisTrpArg 244
 QY 592 ATTAAAAGGCAATGCACTGCTGTGATCGAATGATATAGGAAGAGGCCACACATGAT 651
 Db 245 ValGlnGlyLeuGluAlaArgTrpPheLeuAspAlaTyrAlaArgArgProAspMetAsn 264
 QY 652 CCAGTAGTGTGGAGCTTGCATCTACCTACCTAAATATTGTTCAAGCAAAATTTCAAGAA 711
 Db 265 ProLeuIlePheLysLeuAlaLysLeuAsnPheAsnIleValGlnAlaThrTyrGlnGlu 284
 QY 712 GAGCTCAAGAATCCCTCAGGTGCTGAGAAATACTGGGTTTGTGAGAGCTGCCCTTC 771
 Db 285 GluLeuLysAspIleSerArgTrpAsnSerCysLeuAlaGlnLysLeuProPhe 304
 QY 772 GCAAGGATAGCTGTGGAATCTACTTTTGGATATCTGGGATCATCGAGCCACGTCAG 831
 Db 305 ValArgAspArgIleValGluCysPhePheTrpAlaIleAlaAlaPheGluProHisGln 324
 QY 832 CATGCAAGTCGAAGTAATGATGGGCAAGTCAAGCTCTGATACCTGATACCTGATCAT 891
 Db 325 TyrSerTyrGlnArgLysMetAlaAlaValIleIleThrPheIleThrIleLeuAsp 344
 QY 892 ATTTATGATCTTATGGCACTTGTAGGAACCTCGAACAATTCACCTCACTTCATCGAAGA 951
 Db 345 ValTyrAspValTyrGlyThrIleGluGluLeuLeuLeuLeuLeuLeuLeuLeuLeu 364
 QY 952 TGGGATATAAATCAATCGACCACTTCCGATATACGCAACTGCTGCTTCTTCCACTC 1011
 Db 365 TrpAspAsnLysSerIleSerGlnLeuProTyrTyrMetGlnValCysTyrLeuAlaLeu 384
 QY 1012 AACAACTTCGTGATGATACATCTGATGATGTTATGAGGAGAAAGGCGTCAACCTTATA 1071
 Db 385 TyrAsnPheValSerGluArgAlaTyrAspIleLeuLysAspGluHisPheAsnSerIle 404
 QY 1072 CCTTACTCGGCAATCGTGGTGTGATTTGGCGGATAGTATATGATGAGGACGCGTGG 1131
 Db 405 ProTyrLeuGlnArgSerTrpValSerLeuValGluGlyTyrLeuLysGluAlaTyrTrp 424
 QY 1132 TTCTACGGGGGCAACAACCAAGTTTGGAGAGTATTTGGAGAACTCATGGCAGTCGATA 1191
 Db 425 TyrTyrAsnGlyTyrLysProSerLeuGluGluTyrLeuAsnAlaLysIleSerIle 444
 QY 1192 AGTGGGCTGTATGTTAAGGCACATATCTTCCGAGTAAACAGATTCGTTCAACAGGAG 1251
 Db 445 SerAlaProThrIleSerGlnLeuTyrPheThrLeuAlaAsnSerIleAspGluThr 464
 QY 1252 ACCGTCGACAGTTTGTACAAATACCACGATTTAGTTTCGTTTCGTCATCTCTGTCGG 1311
 Db 465 AlaIleGluSerLeuTyrGlnTyrHisAsnIleLeuTyrLeuSerGlyTyrThrIleLeuArg 484
 QY 1312 CTTCGTGATGATTTGGGACCTCGGTGGAGAGGTGACGACAGGGGAGTGTCCGAAATCA 1371
 Db 485 LeuAlaAspLeuGlyThrSerGlnHisGluLeuGluArgGlyAspValProLysAla 504
 QY 1372 CTTTCAGTGTCTACATGATGATCAATGTCATCGGAGGCGGAGCGCGGAGACACGTCGAAA 1431
 Db 505 IleGlnCysTyrMetAsnAspThrAsnAlaSerGluArgGluAlaValGluHisValLys 524
 QY 1432 TGGCTCATAGCGAGGTGTGGAGAGATGATGATCGGAGAGGGTGTGCGAAGATCTTCCA 1491
 Db 525 PheLeuIleArgGluAlaTrpLysGluMetAsnThrValThrAlaSerAspCysPro 544
 QY 1492 TTCGGCAAGATTTTATGATGTGACGTGATGTTAGTAAGAGATGGCGGAGTTGATGATAC 1551
 Db 545 PheThrAspLeuValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 564
 QY 1552 CATATGGAGATGGGACGCGCACACACACCTTATTATACATCAACAAATGACACAGACC 1611

Db 513 ThrileGlnCysTyrMetIysGluThrAsnAlaSerGluGluAlaValGluHisVal 532
Qy 1429 AAATGGCTCATACGGAGGTGTGGAGAAAGATGAATGGGAGAGGGTGTGGAAGGATCT 1488
Db 533 LysPheLeuIleArgGluAlaTrpIysAspMetAsnThr--AlaIleAlaAlaGlyTyr 551
Qy 1489 CCATTCGGCAAGATTTATAGGATGTGCAGTTGATTAGGAGGATGGCGAGTTGATG 1548
Db 552 PropheProAspGlyMetValIaGlyAlaAlaAsnIleGlyArgValAlaGlnPheIle 571
Qy 1549 TACCATAATGGAGATGGGACGCGCACACACACCCCTATTATATACATCAACAATGACCAGA 1608
Db 572 TyrLeuHisGlyAspGlyPheGlyValGlnHisSerLysThrTyrGluHisIleAlaGly 591
Qy 1609 ACTTATTTCAGCCCTTTGCA 1629
Db 592 LeuLeuPheGluProTyrAla 598

RESULT 13

US-09-900-797-26
; Sequence 26, Application US/09900797
; Publication No. US20030087406A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20030087406A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/900,797
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 598
; TYPE: PRT
; ORGANISM: Salvia officinalis
US-09-900-797-26

Alignment Scores:
Pred. No.: 2,84e-128 Length: 598
Score: 1468.00 Matches: 281
Percent Similarity: 71.12% Conservative: 108
Best Local Similarity: 51.37% Mismatches: 152
Query Match: 50.48% Indels: 6
DB: 10 Gaps: 5

US-09-938-956-6 (1-1632) x US-09-900-797-26 (1-598)

Qy 1 ATCAGAGATCCGGAACATACACCTTCCTGTTGGATGTCAACTTCATCAATCGCTT 60
Db 54 IleArgSerGlyAsnTyrGlnProAlaLeuTrpAspSerAsnTyrIleGlnSerLeu 73
Qy 61 CTCAGTACTATAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGTCTCTTGGTG 120
Db 74 AsnThrProTyrThrGluGluArgHisLeuAspArgIleAlaGluLeuIleValGlnVal 93
Qy 121 AGATGGAATCGAGAAAGAACGGATCAAAATTCGAACTGAGTTGATCGATGACTTG 180
Db 94 ArgIleLeuLeuIysGluIysMetGluProValGlnGlnLeuIleHisAspLeu 113
Qy 181 CAGAGATGGGGTGTCCGATCATTTCCAAATAGTTCAAAGAAATCTTGTCTCTATA 240
Db 114 LysTyrLeuGlyLeuSerAspPhePheGlnAspGluIleGlyValIle 133
Qy 241 TATCTGACCATCATTTTACAGAACCCCTTTCCAAAGAACGAAGGATCTCTACTCC 300

Db 134 TyrAsnGluHisLysCysPheHisAsnAsn--GluValGluLysMetAspLeuTyrPhe 152
Qy 301 ACATCTCTTCATTTAGCTTCCTCAGAGAACATGGTTTTCAGATCGCAGAGAGGTATTC 360
Db 153 ThrAlaLeuGlyPheArgLeuLeuArgGlnHisGlyPheAsnIleSerGlnAspValPhe 172
Qy 361 GATATGTTTCAAGAACAGGAGAGGT--GAGTTCAAAGAAAGCCTTAGCCAGCACACAGA 417
Db 173 AsnCysPheLysAsnGluLysGlyIleAspPheLysAlaSerLeuAlaGlnAspThrLys 192
Qy 418 GGATGTTGCACTGTATGAAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 477
Db 193 GlyMetLeuGlnLeuTyrGluAlaSerPheLeuLeuArgLysGlyGluAspThrLeuGlu 212
Qy 478 TCAGCGAGGAATTCGCCACCAATTTTTCGAGGAAAGTGAACAGAGGGTGTG----- 531
Db 213 LeuAlaArgGluPheAlaThrLysCysLeuGlnLysLeuAspGluGlyGlyAsnGlu 232
Qy 532 GTTGATGGCGACCTTTTAAACAGAATCGCATATCTTTTGGACATCCCTCTTCTTCTTCT 591
Db 233 IleAspGluAsnLeuLeuLeuTrpIleArgHisSerLeuAspLeuProLeuHisTrpArg 252
Qy 592 ATTAAAGGCCAAATGCACCTGTGTGATCGAATGGTATAGGAGAGGCCCGACATGAAT 651
Db 253 IleGlnSerValGluAlaArgTrpPheIleAspAlaTyrAlaArgArgProAspMetAsn 272
Qy 652 CCACTAGTGTGGAGCTTGGCATACTCGACTTAATATATTGTTCAAGCACCAATTTCAAGAA 711
Db 273 ProLeuIlePheGluLeuAlaLysLeuAsnPheAsnIleIleGlnAlaThrHisGlnGln 292
Qy 712 GAGCTCAAGAAATCCTTCAGTGTGGAGAAATACTGGGTTTGTGTGAGAGAGTCCCTTC 771
Db 293 GluLeuLysAspLeuSerArgTrpTrpSerArgLeuCysPheProGluLysLeuProPhe 312
Qy 772 GCAGGATAGACTGTGGGAATGCTACTTTTGAATACTCGGATCATCGACCCACGTCAG 831
Db 313 ValArgAspArgLeuValGluSerPhePheTrpAlaValGlyMetPheGluProHisGln 332
Qy 832 CATCAAGTCAAGGATATGATGGGCAAGTCAACGCTCTGATTACGTTGATCGATGAT 891
Db 333 HisGlyTyrGlnArgLysMetAlaAlaThrIleIleValLeuAlaThrValIleAspAsp 352
Qy 892 ATTTATGATGTCTATGGCACTTAGAAGAACTCGAACAAATTCAGTCACTTCTTCTTCT 951
Db 353 IleTyrAspValTyrGlyThrLeuAspGluLeuGluLeuPheThrAspThrPheLysArg 372
Qy 952 TGGATATAAATCAATCGACCAACTTCCGATTTACATGCAACTGTCTTCTTCTTCTTCT 1011
Db 373 TrpAspThrGluSerIleThrArgLeuProTyrTyrMetGlnLeuCysTyrTrpGlyVal 392
Qy 1012 AACCACTTCGTGCGATGATACATCGTACGATGTTATGAAGAGAAAGCGCTCAACGTTATA 1071
Db 393 HisAsnTyrIleSerAspAlaAlaTyrAspIleLeuLysGluHisGlyPhePheCysLeu 412
Qy 1072 CCTACTCTGGCGCAATCGTGGGTTGATTGGCGGATTAAGTATATGTTAGAGGACCGTGG 1131
Db 413 GlnTyrLeuArgLysSerValValAspLeuValGluAlaTyrPheHisGluAlaLysTrp 432
Qy 1132 TTCTACGGCGGACAAACCAAGTTTCGAAGATTTTGGAGAACTCATCGACGTCGATA 1191
Db 433 TyrHisSerGlyTyrThrProSerLeuAspLeuTyrLeuAsnIleAlaLysIleSerVal 452
Qy 1192 AGTGGCGCTGTATGTAAACGACATATCTTCGAGTAACAGAT--TCGTTCCAAAG 1248
Db 453 AlaSerProAlaIleIleSerProThrTyrPheThrPheAlaAsnAlaSerHisAspThr 472
Qy 1249 GAGACCGTCGACGTTTGTCAATACGAGTTTACTTCGTTGTCATCTCTCTGTTCTG 1308
Db 473 AlaValIleAspSerLeuTyrGlnTyrHisAspIleLeuCysLeuAlaGlyIleIleLeu 492
Qy 1309 CGGCTTGTGATGATTGGGAACCTCGGTGGAAGAGGTGAGCAGAGGGGATGTCCGAAA 1368
Db 493 ArgLeuProAspAspLeuGlyThrSerTyrPheGluLeuAlaArgGlyAspValProLys 512


```
QY 1273 TACCACGATTAGTTGGTTCCTGCTCATCTTCTGCTGCGCTGCTGATGATTTGGGAACC 1332
Db      :      :      :      :      :      :      :      :      :      :
Db 437 AsnProlysinleuGluAlaAenValThrLeuCyArgValileAspAspIleAlaThr 456
QY 1333 TCGGTGAAGAGGTGACGACGAGGGATGTCGCCGAATCACTTCAGTCTACATGATGAC 1392
Db      :      :      :      :      :      :      :      :      :      :
Db 457 TyrGluValGluLysGlyArgGlyGlnIleAlaThrGlyIleGluCytyrMetArgasp 476
QY 1393 TACAATGCATCGGAGGGGAGGCGGACGACGTCGAATGGTGTGATAGCGGAGGTGG 1452
Db      :      :      :      :      :      :      :      :      :      :
Db 477 TyrGlyValSerThrGluLysAlaMetGluLysPheGlnGluMetAlaGluThrAlaTrp 496
QY 1453 AAGAAGATGAATCGGAGAGGGTGTGCAAGGATCTCCATTCGGCAAGATTTATAGGA 1512
Db      :      :      :      :      :      :      :      :      :      :
Db 497 LysAspValAen--GluGlyIleLeuArgProThrProValSerThrGluIleLeuThr 515
QY 1513 TGTGCAGTTGATTAGGAAGAGTGGCGCAGTTGATGATAC---CATATGAGATGGGCAC 1569
Db      :      :      :      :      :      :      :      :      :      :
Db 516 ArgIleLeuAenLeuAlaArgIleleAspValThrTyrLysHisAenGlnAspGlyTyr 535
QY 1570 GGCACACACACACCCCT 1584
Db      :      :      :      :      :      :      :      :      :      :
Db 536 -----ThrHisPro 538
```

RESULT 15

```
US-09-903-012-32
; Sequence 32, Application US/09903012
; Patent No. US20020094557A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094557A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 556
; TYPE: PRF
; ORGANISM: Solanum tuberosum
US-09-903-012-32
```

```
Alignment Scores:
Pred. No.: 4,13e-69 Length: 556
Score: 836.00 Matches: 182
Percent Similarity: 54.68% Conservative: 116
Best Local Similarity: 33.39% Mismatches: 207
Query Match: 28.75% Indels: 40
DB: 9 Gaps: 10
```

US-09-938-956-6 (1-1632) x US-09-903-012-32 (1-556)

```
QY 1 ATGAGACGATCCGGAACTACAACCTTCTCGTTGGGATGTCAACTTCATCCAAATCGCTT 60
Db      :      :      :      :      :      :      :      :      :      :
QY 17 ValArgProIleAlaAspPheSerProSerLeuTrpGlyAspArgPheHisSerPheSer 36
Db      :      :      :      :      :      :      :      :      :      :
QY 61 CTCAGTACTATTAAGGAGGCAACACAGCGATTAGGGCTTCGTAGCTGTCTCTTTGGG 120
Db      :      :      :      :      :      :      :      :      :      :
Db 37 LeuAspAenGlnIleAlaGlyLysTyr-----AlaGlnIleGluThrLeu--- 52
QY 121 AAGATGGAAGTGGAGAAAGAAACCGATCAATT----- 153
Db      :      :      :      :      :      :      :      :      :      :
```

```
Db      :      :      :      :      :      :      :      :      :      :
QY 53 -----LysGluGlnSerArgIleIleLeuSerAlaSerSerArgArgThr 67
154 -----CGACAACCTTGATTCGATCGACTGTCAGAGGATGGGCTGTCGATCATCTTC 207
Db      :      :      :      :      :      :      :      :      :      :
68 LeuAlaGluLysLeuAspLeuIleAspIleValGluArgLeuGlyIleAlaTyHisPhe 87
208 CAAATGAGTTCAAGAAATCTTCTCTATATAT-----CTCGACCAT 252
Db      :      :      :      :      :      :      :      :      :      :
88 GluLysGlnIleAspAspMetLeuAspGlnPheTyrLysAlaAspProAenPheGluAla 107
253 CACTATTACAAGAACCCCTTTTCCAAAAAGAAAGGATCTCTACTCCACATCTCTTTCGA 312
108 HisGluTyr-----AsnAspLeuGlnThrLeuSerValGln 119
313 TTTAGGCTCTCAGAACATGTTTCAAGTCCACAGAGGTATTTCGATAGTTTCAAG 372
120 PheArgLeuLeuArgGlnHisGlyTyrAsnIleSerProLysLeuPheIleArgPheGln 139
373 AACGAGGAGGTGAGTTCAAGAAAGCTTAGCGACGACACAGAGAGATTGTTGCAACTG 432
140 AspAlaLysGlyLysPheLysGluSerLeuCyAsnAspIleLysGlyLeuLeuAenLeu 159
433 TATGAGCTTCCTTCTGTTGACGGAAGGGAACACGCTCGAGTCAGCGAGGGAATTC 492
160 TyrGluAlaSerHisValArgThrHisGlyLysPheLeuGluAlaLeuAlaPhe 179
493 GCCACCAAAATTTTGGAGGAAAAAGTGAACGAGGTGCTGTGATGCGACGCTTTTAAACA 552
180 SerThrAlaHisLeuGluSerAlaAlaProHis-----LeuLysSerProLeuSerLys 197
553 AGAATCGCATATCTTTGGACATCCCTCTTCATTTGGAGGATTAAGGCCAAATGCACCT 612
198 GlnValThrHisAlaLeuGluGlnSerLeuHisLysSerIleProArgValGluThrArg 217
613 GTGTGCGATCGAATGTTATAGGAAGCGCCGACATGAATCCAGTACGATGTTGGAGCTTGC 672
218 TyrPheIleSerIleTyrGluGluGluGlnLysAsnAspValLeuLeuGlnPheAla 237
673 ATCTCGACTTAATATTTTCAAGCACAATTTCAAGAGAGCTCAAGATCTCTTCAGG 732
238 LysLeuAspPheAenLeuLeuGlnMetLeuHisLysGlnGluLeuSerGluValSerArg 257
733 TGGTGGAGAAATCTGGTGTGTTGAGAAGCTGCCCTTCGCAAGGGATAGACTGGTGAA 792
258 TrpTrpLysAspLeuAspPheValThrLeuProTyrAlaArgAspArgAlaValGlu 277
793 TGCTACTTTTGAATCTGGGATCATCGACCGACGTCAGCATGCAAGTGAAGATATG 852
278 CysTyrPheTrpThrMetGlyValTyrAlaGluProGlnTyrSerGlnAlaArgValMet 297
853 ATGGGCAAGTCAACGCTCGATTACGTCGATGATATTTATGATGTCATGGCACC 912
298 LeuAlaLysThrIleAlaMetIleSerIleValAspAspThrPheAspAlaTyrGlyLe 317
913 TTAGAAGAACTCGAACAAATTCATCTGACTTCATTCGAGAGTGGGATATAAATCAATCGAC 972
318 ValLysGluLeuGluIleTyrThrAspAlaIleGlnArgTrpAspIleSerGlnIleAsp 337
973 CAACTTCCGATTTACATGCAACTGTGCTTCTTCGACTCAACACTTCGTCGATGATACA 1032
338 ArgLeuProAspTyrMetLysIleSerTyrLysAlaLeuLeuAspLeuTyrAsnAspTyr 357
1033 TCGTACGATGTTATGAAGAGAAAGGCGTCAAGTATATACCTTACCTCGCGCAATCTGG 1092
358 GluMetGluLeuSerLysAspGlyArgSerAspValValHisTyrAlaLysGluArgMet 377
1093 GTTCATTTGGCGGATAGTATATGTTAGAGCAGCGGTGGTCTTACGGCGGCGCAACCA 1152
378 LysGluIleValArgAsnTyrPheValGluAlaLysTrpPheIleGluGlyTyrMetPro 397
1153 AGTTTGAAGAGTATTTGGAGAACTCATGGCAGTCGATAAGTGGCGCTCTGTATGTAACG 1212
398 ProValSerGluTyrLeuSerAsnAla---LeuAlaThrSerThrTyrTyrLeuLeuThr 416
```

```

QY 1213 CACATATCTTCGCGAGTAACAGATTCTTCACAAAGGAGACCGTCGACAGTTTGTACAAA 1272
Db 417 ThrThrSerTyrLeuGlyMetLysSerAlaAsnLysGlnAspPheGluTrpLeuAlaLys 436
QY 1273 TACCACGATTTAGTTTCGTTGGTGCATCCCTTCCTTCGCGCTTGTGATGATTTGGGAACC 1332
Db 437 AsnProLysIleLeuGluAlaAsnValThrLeuCysArgValIleAspAspIleAlaThr 456
QY 1333 TCGGTGGAGAGGTGAGCAGAGGGGATGTGCCGAATCACTTCAGTCTCATCATGAGTGAC 1392
Db 457 TyrGluValGluLysGlyArgGlyGlnIleAlaThrGlyIleGluCysTyrMetArgAsp 476
QY 1393 TACAATGCATCGAGCGGAGCGCGGAGCAGCACGTGAAATGGCTGATAGCGGAGGTGTGG 1452
Db 477 TyrGlyValSerThrGluLysAlaMetGluLysPheGlnGluMetAlaGluThrAlaTrp 496
QY 1453 AAGAAGATGAATGCGGAGAGGGGTGCGAAGGATTCTCCATTCCGCAAGATTTTATAGGA 1512
Db 497 LysAspValAsn--GluGlyIleLeuArgProThrProValSerThrGluIleLeuThr 515
QY 1513 TGTGCAGTTGATTAGGAGGATGCGCAGTTGATGTAC---CATATGGAGATGGGCAC 1569
Db 516 ArgIleLeuAsnLeuAlaArgIleIleAspValThrTyrLysHisAsnGlnAspGlyTyr 535
QY 1570 GGCACACACACCCCT 1584
Db 536 -----ThrHisPro 538

```

Search completed: February 25, 2004, 06:03:40
Job time : 113 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 25, 2004, 05:40:59 ; Search time 32 Seconds

(without alignments)
9811.536 Million cell updates/sec

Title: US-09-938-956-6

Perfect score: 2908

Sequence: 1 atgagacgacgcgaaacta.....tattcgagcccttgcata 1632

Scoring table:

BLSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 566732

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-O=/cgn2_1/USPTO_epool_P/US0938956/runat_24022004_144701_3078/app_query.fasta_1.1799
-DB=PIR_78 -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=expct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=PCO -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0938956@cgn_1_1_43 @runat_24022004_144701_3078 -NCPU=6 -ICPU=3
-NO_WMAP -LARGUEURY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2876	98.9	599	2 A48863	limonene cyclase -
2	1147	39.4	591	2 H84633	probable limonene
3	1084.5	37.3	600	2 F71434	probable limonene
4	1039	35.7	1024	2 G71434	probable limonene
5	822.5	28.3	559	2 T08174	sequiterpene cycl
6	812	27.9	520	2 A56118	vetispiradiene syn
7	809.5	27.8	550	2 T03714	5-epi-aristolochene
8	796	27.4	554	2 S68366	(+)-delta-cadinene
9	776	26.7	554	2 S68365	germacrene C synth
10	734	25.2	548	2 T06266	germacrene C synth
11	724	24.9	548	2 T06265	hypothetical prote
12	623	21.4	632	2 E96723	probable terpene s
13	600	20.6	608	2 G86443	hypothetical prote
14	565	19.4	421	2 C96642	

15	562.5	19.3	350	2	C56118	vetispiradiene syn
16	552.5	19.0	530	2	G96588	hypothetical prote
17	545.5	18.8	598	2	H86460	hypothetical prote
18	530.5	18.2	573	2	T05328	hypothetical prote
19	515.5	17.7	300	2	B56118	vetispiradiene syn
20	510.5	17.6	598	2	F96684	probable terpene s
21	508	17.5	535	2	T06285	hypothetical prote
22	498.5	17.1	582	2	C71424	hypothetical prote
23	484	16.6	471	2	T06287	hypothetical prote
24	479	16.5	598	2	T05009	probable vetispira
25	457	15.7	612	2	T05331	hypothetical prote
26	429.5	14.8	383	2	D71424	hypothetical prote
27	414	14.2	785	2	G96825	hypothetical prote
28	414	14.2	785	2	T52059	ent-kaurene syntha
29	405	13.9	789	2	T05329	hypothetical prote
30	359	12.3	789	2	T05672	ent-kaurene syntha
31	353.5	12.2	870	2	A96637	hypothetical prote
32	309.5	10.6	802	2	D85035	ent-kaurene synthe
33	275	9.5	823	2	T02959	kaurene synthase A
34	272.5	9.4	801	2	T06783	ent-kaurene syntha
35	138.5	4.8	1169	2	S38181	floculation prote
36	138	4.7	203	2	H96525	probable terpene c
37	137.5	4.7	148	2	T03982	5-epi-aristolochene
38	130.5	4.5	2829	2	A42771	reticulocyte-bind
39	130	4.5	850	2	G70332	conserved hypotet
40	129.5	4.5	534	2	T39903	serine-rich protei
41	127	4.4	686	2	T10684	hypothetical prote
42	125.5	4.3	1030	2	S73944	hypothetical prote
43	124	4.3	1939	2	T48175	myosin heavy chain
44	123	4.2	1713	2	A55347	adhesive ligand ep
45	123	4.2	2672	2	A48126	translation activa

ALIGNMENTS

RESULT 1

A48863

limonene cyclase - spearmint

C/Species: Mentha spicata (spearmint)

C/Date: 12-May-1995 #sequence_revision 19-May-1995 #text_change 05-Nov-1999

C/Accession: A48863

J. Biol. Chem. 268, 23016-23024, 1993

A/Title: 4S-limonene synthase from the oil glands of spearmint (Mentha spicata). cDNA i

A/Reference number: A48863; MUID:94043077; PMID:8226816

A/Accession: A48863

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-599 <COL>

A/Cross-references: GB:113459; NID:9410229; PIDN:AAC37366.1; PID:9410230

C/Superfamily: vetispiradiene synthase 1

Alignment Scores:
Pred. No.: 4.87e-204 Length: 599
Score: 2876.00 Matches: 542
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.90% Indels: 0
DB: 2 Gaps: 0

US-09-938-956-6 (1-1632) x A48863 (1-599)

Qy	4	AGACGATCGGAAACTACACCTTCTGTTGGATGTCACTTCATTCATTCGCTTC	63
Db	58	AGAGGserGlyAsnTrpAsnProserArgTrpAspValAsnPhellegInSerLeuLeu	77
Qy	64	ATGACTATAGAGGACAAACACGTGATTAGGGTTCTTACGCTGCTCATCTTTGGTGAAG	123
Db	78	SerAspTyLysGluAspLysHisValIleAsgAlaSerGluLeuValThrLeuValLys	97
Qy	124	ATGGACTCGGAGAGAACCGATCAATTCGACACTGACTGATTCATTCGACTGAC	183
Db	98	MetGluLeuGluLysGluThrAspGlnIleArgGlnLeuGluLeuLeuAspLeuGln	117

184 AGGATGGGGCTCGCATCATTTTCAAAATAGATTCAAGAAATCTTCTCTATATAT 243
 118 ArgMetGlyLeuSerAspHisPheGlnAsnGluPheLysGluLeuSerSerIleTyr 137
 244 CTCGACCATCATTTACAGAACCTTTTCCAAAGAGAGAGAGGATCTCTACTCCACA 303
 138 LeuAspHisHisTyrTyrLysAsnProPheProLysGluGluArgAspLeuTyrSerThr 157
 304 TCTCTTGCATTTAGGCTCCTCAGAAACATGCTTTTCAAGTCGCAACAAGAGGATTCGAT 363
 158 SerLeuAlaPheArgLeuLeuArgGluHisGlyPheGlnValAlaGlnGluValPheAsp 177
 364 AGTTTCAAGAACGAGAGGGTGAAGTTCAAGAAAGCCTTAGCGACGACACAGAGGATTG 423
 178 SerPheLysAsnGluGluGluPheLysGluSerLeuSerAspAspThrArgGlyLeu 197
 424 TTGCAACTCTATCAAGCTTCTCTTCTGTTGACGAGGCGAAACACACCTCGAGTCAGCG 483
 198 LeuGlnLeuTyrGluAlaSerPheLeuLeuThrGluGluGluThrLeuGluSerAla 217
 484 AGGAAATTCGCCACCAATTTTGGAGGAAAGTGAACGAGGGTGGTGTGATGGCGAC 543
 218 ArgGluPheAlaThrLysPheLeuGluGluLysValAsnGluGlyGlyValAspGlyAsp 237
 544 CTTTAAACAAGATCGCATATTTCTTGGACATCCCTCTTATTGGAGGATTAAGAGCCA 603
 238 LeuLeuThrArgIleAlaTyrSerLeuAspIlePheLeuHisTyrPheIleLysArgPro 257
 604 AATGCACCTGTGTGATCGAATGGTATAGGAAGCGCCGACATCAATCCAGTAGTGTG 663
 258 AsnAlaProValTyrIleGluTyrPheArgLysArgProAspMetAspProValValLeu 277
 664 GAGCTGGCCATCTGACATTAATATGTTCAAGCACATTTCAAGAGAGAGCTCAAGAA 723
 278 GluLeuAlaIleLeuAspLeuAsnIleValGlnAlaGlnPheGlnGluLeuLysGlu 297
 724 TCTTCAGGTGGTGGAGAATACTCGGTTTGTGTGAGAGCTCCCTTCGCAAGGGATAGA 783
 298 SerPheArgTyrPheArgAsnThrGlyPheValGluLysLeuPheAlaArgAspArg 317
 784 CTGTTGGAATGCTACTTTGGATCTGGATCATCGGACCGCCAGTCAGCATGCA 843
 318 LeuValGluCysTyrPheTyrPheThrAsnThrGlyIleLeuGluProArgGlnHisAlaSerAla 337
 844 AGGATAATGATGGCAAGCTCAACGCTCTGATGATCGATGATATTTATGATGTC 903
 338 ArgIleMetMetGlyLysValAsnAlaLeuIleThrValIleAspAspIleTyrAspVal 357
 904 TATGGCACCTTAGAAGACTCGAACAATTCATCTGACCTCATTCGAGATGGGATATAAC 963
 358 TyrGlyThrLeuGluGluLeuGluGlnPheThrAspLeuIleArgArgTyrAspIleAsn 377
 964 TCAATCGACCACTTCCGATTTACATGCAACTGTGTTTCTTGTGCACTCAACAACCTCGTC 1023
 378 SerIleAspGlnLeuProAspTyrMetGlnLeuCysPheLeuAlaLeuAsnAsnPheVal 397
 1024 GATGATACATCGTACGATGTTATAGAGGAGAGCGGTCAACGTTATACCTTACCTGCGG 1083
 398 AspAspThrSerTyrAspValMetLysGluLysGlyValAsnValIleProTyrLeuArg 417
 1084 CAATCGTGGTGTGATTTGGCGGATAGTATATGTTAGAGCAGCGGTGTTCTACCGCGG 1143
 418 GlnSerTyrValAspLeuAlaAspLysTyrMetValGluAlaArgTyrPheTyrGlyGly 437
 1144 CACAAACCAAGTTTGAAGAGTATTTGGAGAACTCATGGCAGTCGATAGTGGCCCTGT 1203
 438 HisLysProSerLeuGluGluTyrLeuGluAsnSerTyrGlnSerIleSerGlyProCys 457
 1204 ATGTTAACGCATATTTCCGAGTAAACAGATTCGTTTCAAGAGGAGCGCTCGACAGT 1263
 458 MetLeuThrHisIlePhePheArgValThrAspSerPheThrLysGluThrValAspSer 477

1264 TTGTACAAATACCAACGATTTAGTTTGGTTCATCCTTGGTTCGGGCTTGTGATGAT 1323
 478 LeuTyrLysTyrHisAspLeuValArgTyrPheSerPheValLeuArgLeuAlaAspAsp 497
 1324 TTGGGAACCTCGGTGCAAGAGGTGAGCAGAGGGGATGTCCGAAATCACTTCAAGTCGAC 1383
 498 LeuGlyThrSerValGluGluValSerArgGlyAspValProLysSerLeuGlnCysTyr 517
 1384 ATGAGTGACTCAATCATCGATCGAGGCGGAGCGCGAAGACACGTGAAATGGCTGATAGCG 1443
 518 MetSerAspTyrAsnAlaSerGluAlaGluAlaArgLysHisValLysTyrLeuIleAla 537
 1444 GAGGTGTGGAAGAGATGAATCGGAGAGGGTGTGCAAGGATTCCTCATTCGCAAAAGAT 1503
 538 GluValTyrLysLysMetAsnAlaGluArgValSerLysAspSerProPheGlyLysAsp 557
 1504 TTTATAGGATGTCTGATGTTAGTGAAGATGCGCAGTGTGATGATACATTAATGAGAT 1563
 558 PheIleGlyCysAlaValAspLeuGlyArgMetAlaGlnLeuMetTyrHisAsnGlyAsp 577
 1564 GGCACGGCACCAACACCTTATATACATCAACAAATGACCAAGAACCTTATTCGAGGCC 1623
 578 GlyHisGlyThrGlnHisProIleIleHisGlnGlnMetThrArgThrLeuPheGluPro 597
 1624 TTTGCA 1629
 598 PheAla 599

RESULT 2
 H84633
 probable limonene cyclase [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C;Accession: H84633
 R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Mofatt, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, I.
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
 Nature 402, 761-768, 1999
 A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A;Reference number: AB4420; PMID:20083487; PMID:10617197
 A;Accession: H84633
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-591 <STO>
 A;Cross-references: GB:AE002093; NID:G4115381; PIDN:AAD03382.1; GSPDB:GN00139
 C;Genetics:
 A;Map position: 2
 C;Superfamily: vetispiradiene synthase 1

Alignment Scores:
 Pred. No.: 1,43e-76 Length: 591
 Score: 1147.00 Matches: 239
 Percent Similarity: 64.97% Conservative: 119
 Best Local Similarity: 43.38% Mismatches: 173
 Query Match: 39.44% Indels: 20
 DB: 2 Caps: 10

US-09-938-956-6 (1-1632) x H84633 (1-591)

Qy 4 AGACGATCCGAAACTCAACCTTCTCGTTGGGATGTCACATTCATCCATCGCTTCTC 63
 Db 46 ArgArgSerAlaAsnTyrGlnProSerArgTyrAspHisHisLeuLeuSerValGlu 65
 Qy 64 AGTCACTATAGAGGAGCAACACGATGATAGGCTTCTGAGCTGTC-----ACTTTG 117
 Db 66 AsnLysPheAlaLysAspLysArgVal---ArgLysArgAspLeuLeuLysGluLysVal 84
 Qy 118 GTGAGGATGCACTGGAGAAAGAAACGATCAATTCGACCACTTGAATGATGATGAC 177
 Db 85 ArgLysMetLeuAsnAspGluGlnLysThrTyrLeuAspGlnLeuGluPheLeuAsp 104
 Qy 178 TTGCAGAGGATGGGGCTGTCGGATCATTTCCAAATAGATTCAAGAAATCTTGTCTCTCT 237

Db	457	LeuSerIleGlnValLeuGluThrLeuSerGlnHisGlnGlnAsnValValArgCysSer	476
QY	1297	TCCTTCCTCTCGGGCTTCTGATGATTTGGGAACCTCGGTGGAGAGGTGACGACAGGG	1356
Db	477	SerSerValPheArgLeuAlaAsnAspLeuValThrSerProAspGluLeuAlaArgGly	496
QY	1357	GATGTGCGGAAATCACTTCAGTCTCATCATGAGTCATACATGCATCGAGGCGGAGCG	1416
Db	497	AspValCysIysSerIleGlnCysTyrMetSerGluThrGlyAlaSerGluAspLysAla	516
QY	1417	CGGAAGCACGTGAATGGCTGATAGCGGAGGTGTGGAGAGATGAATCGGAGAGGGTG	1476
Db	517	ArgSerHisValArgGlnMetIleAsnAspLeuThrPaspGluMetAsnTyrGluLysMet	536
QY	1477	TCGAAGGATTTCTCCA---TTCGGCAAGATTTTATATAGGATGTCAGTTGATTTAGGAAG	1533
Db	537	AlaHisSerSerIleLeuHisAspPheMetGluThrValIleAsnLeuAlaArg	556
QY	1534	ATGGCGCAGTTGATGATACCATATATGAGATGGCAGCGGCACA---CAACACCCCTATTATA	1590
Db	557	MetSerGlnCysMetTyrGlnTyrGlyAspGlyHisGlySerProGluLysAlaLysIle	576
QY	1591	CATCAACAATGACCAAGACCTTATTTCGAGCCC	1623
Db	577	ValAspArgValMetSerLeuLeuPheAsnPro	587
RESULT 3			
F71434			
probable limonene cyclase - Arabidopsis thaliana			
C:Species: Arabidopsis thaliana (mouse-ear cress)			
A:Variety: Columbia			
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 20-Jun-2000			
C:Accession: F71434			
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, A.; vanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Fyfe, Nature 391, 485-488, 1998			
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puijchert, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, C.; Chwalatys, N.			
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis			
A:Reference number: A71400; MUID:98121113; PMID:9461215			
A:Accession: F71434			
A:Status: preliminary; nucleic acid sequence not shown; translation not shown			
A:Molecule type: DNA			
A:Residues: 1-600 <BEV>			
A:Cross-references: GB:Z97341; NID:92244991; PID:92245028			
C:Genetics:			
A:Map position: 4C0P9-4G3845			
C:Superfamily: vetispiradiene synthase 1			
Alignment Scores:			
Pred. No.:	5	8e-72	Length: 600
Score:	1084.50	Matches:	218
Percent Similarity:	63.34%	Conservative:	1.2
Best Local Similarity:	41.84%	Mismatches:	158
Query Match:	37.29%	Indels:	33
DB:	2	Gaps:	8
US-09-938-956-6 (1-1632) x F71434 (1-600)			
QY	103	GAGCTGGTCACCTTGGTGAAGATGGAACTGGAGAAAGAAACGGATCAAAATTCGACAACTT	162
Db	96	GluValGluThrAlaGluLysAlaIleLeuPheLysGlu-----GluLeu	110
QY	163	GAGTTGATCGATGACTTCGAGAGGATGGCGTCCGATCATTTCCAAAATGAGTTCAA	222
Db	111	GluMetIleAspSerLeuGlnArgLeuGlyIleSerTyrHisTyrLysHisGluIleHis	130
QY	223	GAATCTTGCTCTATATATCTCGACCATCATCTATTACAGAAACCCCTTTTCCAAAGAA	282
Db	131	AspIleLeuArgLysIleHis---AspGlnHis-----GlyLulIleGluArgGlu	146

Qy	283	GA	AAGG	ATCTCTACTCC	ATCATCTCTTTGCA	TTTAGGCTCTCTCAGAGAACATGTTTTC	CAA	342															
Db	147	Thr	Gln	Asp	Leu	His	Ala	Thr	Ser	Leu	Glu	Phe	Ile	Leu	Leu	Arg	Gln	His	Gly	Phe	Asp	166	
Qy	343	GT	CG	CA	CAC	AAG	GGT	TAT	TCG	ATAGTTT	CA	GA	GA	CG	AGG	GGT	GAGTT	CA	AA	GA	AA	GCCTT	402
Db	167	Val	Ser	Gln	Asp	Ala	Phe	Asp	Val	Phe	Ile	Ser	Glu	Thr	Gly	Glu	Phe	Arg	Lys	Thr	Leu	186	
Qy	403	AG	CG	AC	CAC	ACC	AGG	ATGTT	TC	CAACTGT	ATGAAGCTTC	TCTT	CTT	CTT	CTT	CTT	CTT	CTT	CTT	CTT	CTT	CTT	462
Db	187	His	Ser	Asp	Ile	Lys	Gly	Leu	Leu	Ser	Leu	Tyr	Glu	Ala	Ser	Tyr	Phe	Ser	Met	Asp	Ser	206	
Qy	463	GA	RA	CC	AC	CGT	CG	AGT	CAG	CGG	GAATTC	CG	CA	CA	CA	AA	TTT	TG	GA	GA	AA	GA	522
Db	207	Glu	Phe	Lys	Leu	Lys	Glu	Thr	Arg	Gly	Ile	Ala	Asn	Lys	Arg	Leu	Ser	Glu	Phe	Val	Ala	226	
Qy	523	GAG	GGT	GGT	GT	-----	GAT	GGC	AGC	CTTTT	TA	CA	GA	AT	CG	CAT	AT	564					
Db	227	Glu	Ser	Ser	Lys	Thr	Ile	Cys	Arg	Glu	Asp	Glu	Thr	Tyr	Ile	Leu	Glu	Met	Val	Lys	Arg	246	
Qy	565	TC	TTT	G	A	C	A	T	C	CTT	TC	AT	TG	G	A	GG	CA	AT	T	A	A	AG	624
Db	247	Ala	Leu	Glu	Thr	Pro	Tyr	His	Trp	Ser	Ile	Arg	Leu	Glu	Ala	Arg	Tyr	Ile	Asn	266			
Qy	625	TG	T	AT	AG	GA	AG	CG	CC	G	CA	T	CA	A	T	C	T	AG	T	GT	TG	AG	684
Db	267	Val	Tyr	Glu	Lys	Lys	His	Glu	Met	Asn	Pro	Leu	Leu	Leu	Glu	Phe	Ala	Ile	Asp	Phe	286		
Qy	685	A	T	A	T	G	T	T	CA	AG	C	A	A	T	T	CA	A	G	A	G	A	G	744
Db	287	Asn	Met	Leu	Gln	Ala	Asn	His	Gln	Glu	Glu	Leu	Lys	-----	Leu	Ile	Ser	Ser	302				
Qy	745	ACT	GGG	TTT	GT	TG	AG	AG	CTCC	CTT	CG	CA	AGG	AT	AG	CT	TG	TG	GA	T	C	T	804
Db	303	Thr	Gly	Leu	Met	Lys	Gln	Leu	Asp	Phe	Val	Arg	Asp	Arg	Ile	Thr	Glu	Ser	Tyr	Phe	Trp	322	
Qy	805	A	A	T	A	C	T	CG	G	A	T	C	A	G	C	A	G	T	C	A	A	G	864
Db	323	Thr	Ile	Gly	Ile	Phe	Tyr	Glu	Pro	Glu	Phe	Lys	Tyr	Cys	Arg	Lys	Ile	Leu	Thr	Lys	Ile	342	
Qy	865	A	A	C	C	T	G	A	T	CG	G	T	A	G	A	T	A	T	T	A	T	A	924
Db	343	Phe	Met	Leu	Ile	Val	Ile	Met	Asp	Asp	Ile	Tyr	Asp	Ile	Tyr	Gly	Thr	Leu	Glu	Glu	Leu	362	
Qy	925	GA	CA	A	T	T	C	A	T	CG	A	G	A	T	G	G	A	T	A	A	A	C	984
Db	363	Glu	Leu	Phe	Thr	Asn	Val	Val	Glu	Lys	Trp	Asp	Val	Asn	His	Val	Glu	Arg	Leu	Pro	Asn	382	
Qy	985	T	A	C	A	T	G	C	A	T	G	C	A	C	A	A	C	T	T	C	G	A	1044
Db	383	Tyr	Met	Arg	Met	Cys	Phe	Leu	Phe	Leu	Tyr	Asn	Glu	Ile	Asn	Gln	Ile	Gly	Tyr	Asp	Val	402	
Qy	1045	A	T	G	A	G	A	G	A	A	G	C	G	T	A	T	A	C	C	T	A	C	1104
Db	403	Leu	Arg	Ser	Lys	Gly	Leu	Asn	Val	Ile	Pro	Tyr	Leu	Lys	Gln	Val	Trp	Thr	Asp	Leu	Phe	422	

Db	503	MetAlaArgGlyAspSerProLysSerValGlnCysTyrMetTyrGluThrArgAlaSer	522
Qy	1405	GAGCGCGAGCGCGGAGACACGTCGAATGCGTCATACGCGAGGTGTGGAAGAAAGATGAAT	1464
Db	523	GluGluGluAlaArgArgHisMetGlnSerMetIleSerAspSerTrpAspIleIleAsn	542
Qy	1465	CGCGAG---AGGGTGTCTCAGAGGATTCCTCATTCGGCAACGATTATTATAGATGTGGCAGTT	1521
Db	543	SerAspLeuIleThrAlaHisThrSerSerLeuProArgGlyPheLeuAlaAlaAla	562
Qy	1522	GATTATGAAGAGTATGGCGAGTGTGATGTACATTAATGCAGATGGCAGCGACACACACAC	1581
Db	563	AsnLeuAsnArgValValGlnCysIleTyrArgHisGlyAspGlyHisGlySer-----	580
Qy	1582	CCATTATATCATCAACAAATGCAGCAACCC-----TTATTCGAG	1620
Db	581	-----ProGluLysThrLysThrValAspTyrIleGlnSerValLeuPheAsn	596
Qy	1621	CCC	
Db	597	Pro 597	
RESULT 4			
G71434			
probable limonene cyclase - Arabidopsis thaliana			
C;Species: Arabidopsis thaliana (mouse-ear cress)			
A;Variety: Columbia			
C;Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998			
R;Accession: G71434			
R;By: Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, A.; vanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; F. P. Nature 391, 485-488, 1998			
A;Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Pui-erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, C.; Chaiwatzi, N.			
A;Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabido			
A;Reference number: A71400; MUID:98121113; PMID:9461215			
A;Accession: G71434			
A;Status: preliminary; nucleic acid sequence not shown; translation not shown			
A;Molecule type: DNA			
A;Residues: 1-1024 <BEV>			
A;Cross-references: GB:297341; NID:g2244991; PID:e327011; PID:g2245029			
C;Genetics:			
A;Map position: 4COP9-4G3845			
Alignment Scores:			
Pred. No.:			
Score:			
Percent Similarity:			
Best Local Similarity:			
Query Match:			
Dp:			

```

QY 295 TACTCCATCTCTTGCTAGCTCCTCAGAGCAATGCTTTTCAAGTCGCACAGAG 354
Db 78 TyrAlaThrAlaLeuGluPheArgLeuLeuArgGlnHis-----Asp 91
QY 355 GTATTGATGATTTCAAGACAGAGGGTGAAGTTCAAGAAACCTTAGCGAC---GAC 411
Db 92 ValPheAsp-----GlyAsnIleGlyValAspLeuAspAspLysAsp 105
QY 412 ACCAGAGATTGTTGCAACTGATGAAGCTTCTCTGTTGCGGAGCGGAACCAAG 471
Db 106 IleLysGlyIleLeuSerLeuTyThrLysArgSerLeuSerThrArgIleAspThrLys 125
QY 472 CTCGAGTCAGCGAGGGAATCCGCCACCAAAATTTTGGAGGAA-----AAAGTGAACGAG 525
Db 126 LeuLysGluSerIleTyThrLysArgLeuArgLysPheValGluValAsnLys 145
QY 526 GGTGGTGTGATGGCGACCTTTAAACAAGA---ATCCGATATCTTTGGACATCCCTCTT 582
Db 146 AsnGluThrLysSerTyThrLeuArgArgMetValIleHisAlaLeuGluMetProTy 165
QY 583 CATTTGGAGGATTAAGAGCCAAATGCGACCTGTGGATCGAATGCTATAGGAAGGCC 642
Db 166 HisArgArgValGlyArgLeuGluAlaArgTrpTyRileGluValTyRileGluArgHis 185
QY 643 GACATGAATCAGTAGTGTGGAGCTTGGCATACTCGACTTAATAATTTCAAGCACAA 702
Db 186 AspMetAsnProIleLeuLeuGluLeuAlaLysLeuAspPheAsnPheValGlnAlaIle 205
QY 703 TTTCAAGAGAGCTCAAGAATCCTTCAGTGTGGTGGAGAAATACTGGTGTGTGAGAAG 762
Db 206 HisGlnAspGluLeuLysSerLeuSerSerTrpSerLysThrGlyLeuThrLysHis 225
QY 763 CTGCCCTTCGCAAGGATAGACTGGTGAATGCTACTTTTGGAAATCTGGGATCATCGAG 822
Db 226 LeuAspPheValArgAspArgIleThrGluGlyTyR-PheSerSerValGlyValMetTy 245
QY 823 CCAGTCGACATGCAAGTCAAGGATAATGATGGGCAAGTCAACGCTCTGATTAGGTG 882
Db 246 GluProGluPheAlaTyHisArgGlnMetLeuThrLysValPheMetLeuIleThr 265
QY 883 ATCGATGATATTTATGATGCTATGGCACCTTGAAGAACTCGAACAATTCACCTGACCTC 942
Db 266 IleAspAspIleTyRAspIleTyRLeuGluGluLeuGlnLeuPheThrIle 285
QY 943 ATTGGAAGATGGGATATAACTCAATCGACCAACTTCCGATTAATGCAACGTGCTTT 1002
Db 286 ValGluLysTrpAspValAsnArgLeuGluGluLeuProAsnTyRMetLysLeuCysPhe 305
QY 1003 CTTCGACTCAACAACTTCGTTCGATATACATCGTATGATGAAGGAGAAAGCGTC 1062
Db 306 LeuCysLeuValAsnGluIleAsnGlnIleGlyTyR-PheValLeuArgAspLysGlyPhe 325
QY 1063 AAGTTATACCTTACTCGCGCAATCGTGGGTGATTGGCGGATGAATATATGTPAGAG 1122
Db 326 AsnValIleProTyRLeuLysGluSerTrpAlaAspMetCysThrThrPheLeuLysGlu 345
QY 1123 GCAGGTGTCTACGCGCGGCACAAACCAAGTTTGGAGAGTATTGGAGACTCATGG 1182
Db 346 AlaLysTrpTyRLeuLysGlyTyRLeuLysProAsnPheGluGluTyRMetGlnAsnGlyTrp 365
QY 1183 CAGTCGATAAGTGGCGCTGTATGTTAAACGCACATATTTCTCGAGTACACAGATTCGTTTC 1242
Db 366 IleSerSerValProThrIleLeuLeuHisLeuPheCysLeuLeuSerAsp----- 383
QY 1243 ACAAGAGACCTCGACAGCTTTGTACAATAC---CAGATTATTGTTGGTGCATCC 1299
Db 384 -----GlnThrLeuAspIleLeuGlySerTyRAsnHisSerValValArgSerAla 401
QY 1300 TTGCTTCTGGGCTGTGCTGATGTTTGGGAACCTCG----- 1335
Db 402 ThrIleLeuArgLeuAlaAsnAspLeuAlaThrSerValSerHisGlyPheThrThr 421
QY 1336 -----GTGGAAGAGGTGACGAGGGGATGTGCGGAATTCATCTTCAGTGTACATGAGT 1389

```

```

Db 422 TyrAsnThrGluGluLeuAlaArgGlyAspThrMetLysSerValGlnCysHis 441
QY 1390 GACTACAATGCTATCGGAGCGCGAGGCGGAAAGCACGTGCAATGCTGATACGCGAGGTG 1449
Db 442 GluThrGlyAlaSerGluAlaGluSerArgAlaTyRileGlnGlyIleIleGlyValAla 461
QY 1450 TGGAGAGAGATCAATGCGGAGAGGGTGTGGAAGGATTCTCCATTGCGGCAAGATTTATA 1509
Db 462 TrpAspAspLeuAsnMetGluLys-----LysSerCysArgLeuHisGlnGlyPheLeu 479
QY 1510 GGATGTGCGAGTTCATTAGGAAGGATGCGCAGTGTGATCCATATCGACATGGGAC 1569
Db 480 GluAlaAlaAlaAsnLeuGlyArgValAlaGlnCysValTyRLeuTyRLeuGlyAspGlyHis 499
QY 1570 GGC 1572
Db 500 Gly 500

RESULT 5
T08174
sesquiterpene cyclase (EC 2.5.1.1) - pepper
C:Species: Capsicum annuum (pepper)
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: T08174
R:Back, K.; Shin, D.H.; He, S.
A:Title: Cloning and bacterial expression of sesquiterpene cyclase, a key branch point
A:Reference number: Z16395; MUID:99033462; PMID:9818674
A:Accession: T08174
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-559 <BAC>
A:Cross-references: EMBL:AF061285; NID:g3108342; PIDN:AAC61260.1; PID:g3108343
A:Function:
A:Pathway: the synthesis of phytoalexin capsidiol
A>Note: induced by UV
C:Superfamily: vetispiradiene synthase 1
C:Keywords: isoprenoid biosynthesis; transferase

Alignment Scores:
Pred. No.: 1,186-52 Length: 559
Score: 822.50 Matches: 182
Percent Similarity: 52.85% Conservative: 115
Best Local Similarity: 32.38% Mismatches: 222
Query Match: 28.28% Indels: 43
DB: 2 Gaps: 9

US-09-938-956-6 (1-1632) x T08174 (1-559)
QY 1 ATGAGACGATCGGAAACTACAAACCTTCTCGTGGGATGCAACTTCATCCAAATCGCTT 60
Db 19 IleArgProValAlaAlaAspPheSerProSerLeuTrpGlyAspArgPheLeuSerPheSer 38
QY 61 CTCAGTGAATTAAGGAGGACAAACACGTGATTAGGGCTTCTGAGCTGGTCACTTGGTG 120
Db 39 IleAspAsnGlnValGluThrLysTyRala----- 48
QY 121 AAGATGGAACCTGGAGAAAGAAACCGATCAAAATTCGACAA----- 159
Db 49 ---GlnGluIleGluProLeuLysGluGlnThrArgSerMetLeuLeuAlaSerGlyArg 67
QY 150 -----CTTGAGTTCATCGATCACTTCAGAGGATGGGGCTGTCCTCCATCAT 204
Db 68 LysLeuSerGluThrLeuAsnLeuIleAspValIleGluArgLeuGlyIleAlaTyRHis 87
QY 205 TTCCAAATGAGTTCAAAGAAATCTTCTCTATATATCTCGACCATCACTATTACAAG 264
Db 88 PheGluLysGluIleAspGluIleLeuAspArgIleTyRAsnGluAsnSerAsnPheGlu 107
QY 265 AACCTTTTCCAAAGAAAGAGGAGTCTCTACCCACATCTCTTGGATTTAGGCTCTC 324
Db 108 GlyAspValTyRAsnGlu-----AspLeuCysThrCysArgLeuGlnPheArgLeu 125

```


Db 145 valGlyHisLeuGluSerAlaAlaProHis-----LeuLysSerProLeuSerLysGln 162
 QY 556 ATGCGATATCTTTGGACATCCCTCTTCATGAGGATTAAAGGCCAAATGCACCTGTG 615
 Db 163 ValThrHisAlaLeuGlnSerLeuHisLysSerIleProArgValGluIleArgTyr 182
 QY 616 TGGATCGAATGCTATAGGAAGGCCCGACATGAATCCAGTAGTGTGGAGTTCGCATA 675
 Db 183 PheIleSerIleTyrGluGluGluGluPheLysAsnAspLeuLeuLeuArgPheAlaLys 202
 QY 676 CTCGACTTAATATCTTCAAGCACAAATTTCAAGAGAGCTCAAGAGATCTTCAGGTGG 735
 Db 203 LeuAspTyrAsnLeuLeuGlnMetLeuHisLysHisGluLeuSerGluValSerArgTyr 222
 QY 736 TGGAGAAATCTGGGTTTGTGAGAGCTGCCCTTCGCAAGGATAGACTGGTGAATGC 795
 Db 223 TrpLysAspLeuAspPheValThrThrLeuProTyrAlaArgAspAlaValGluCys 242
 QY 796 TACTTTTGGAAATCTGGGATCATCGACCCACGCTCAGCATGCAAGTGCAGGATATGATG 855
 Db 243 TyrPheTrpThrMetGlyValTyrAlaGluProGlnTyrSerGlnAlaArgValMetLeu 262
 QY 856 GCGAAAGTCACGCTCTGATTACGGTGATCGATGATATTTATGATGCTATGGCACCTTA 915
 Db 263 AlaLysThrIleAlaMetIleSerIleValAspAspThrPheAspAlaTyrGlyIleVal 282
 QY 916 GAAGAACTCGAACATTTCACTACCTCATTCGAAGATGGATATAAATCAATCGACCAA 975
 Db 283 LysGluLeuGluValTyrThrAspAlaIleGlnArgTrpAspIleSerGlnIleAspArg 302
 QY 976 CTTCCCGATATCATCGAATGCTGCTTTCTTGCACTCAACAACTTCGTGATGATACATCG 1035
 Db 303 LeuProGluTyrMetLysIleSerTyrLysAlaLeuLeuAspLeuTyrAspTyrGlu 322
 QY 1036 TAGATGTTATGAAGGAAGGCGTCAAGTTATACCTACCTCGCCAAATCGTGCGTT 1095
 Db 323 LysGluLeuSerLysAspGlyArgSerAspValValHisTyrAlaLysGluArgMetLys 342
 QY 1096 GATTTGGCGGATAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1155
 Db 343 GluIleValGlyAsnTyrPheIleGluGlyLysTrpPheIleGluGlyTyrMetProSer 362
 QY 1156 TTGGAAGATATTTCGAGAACTCATGCGAGTCGATGATGATGATGATGATGATGATGATG 1215
 Db 363 ValSerGluTyrLeuSerAsnAla---LeuAlaThrSerThrTyrTyrLeuLeuThr 381
 QY 1216 ATATTCTCCGAGTAACAGATTCGTTCAAGAGGACCGTCGACAGTTTGTCAAAATAC 1275
 Db 382 ThrSerTyrLeuGlyMetLysSerAlaThrLysGluHisPheGluTrpLeuAlaThrAsn 401
 QY 1276 CACGATTTAGTTGCTGCTCATCTTCTGCTGCGGCTTGTGATGATGATGATGATGATGATG 1335
 Db 402 ProLysIleLeuGluAlaAsnAlaThrLeuCysArgValValAspAspIleAlaThrTyr 421
 QY 1336 GTGGAAGAGGTGACGAGGGGATGTGCGGAAATCATCTTACGTCTCATGATGATGATGATG 1395
 Db 422 GluValGluLysGlyArgGlyGlnIleAlaThrGlyIleGluCysTyrMetArgAspTyr 441
 QY 1396 AATGCATCGGAGCGGAGCGCGGACGACGTCGAATGCTGATGATGATGATGATGATGATG 1455
 Db 442 GlyValSerThrGluValAlaMetGluLysPheGlnGluMetAlaAspIleAlaTrpLys 461
 QY 1456 AAGTGAATCGGAGAGGGTGTGCAAGATTTCTCCATTCGCGCAAGATTTTATAGGATGT 1515
 Db 462 AspValAsn---GluGluIleLeuArgProThrProValSerSerGluIleLeuThrArg 480
 QY 1516 GCAGTGTATTAGGAAGATGGCGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1572
 Db 481 IleLeuAsnLeuAlaArgIleIleAspValThrTyrLysHisAsnGlnAspGlyTyr--- 499
 QY 1573 ACACAAACCCCT 1584
 Db 500 ---ThrHisPro 502

RESULT 7

T03714
 5-epi-aristolochene synthase - common tobacco
 N:Alternate names: sesquiterpene cyclase
 C:Species: Nicotiana tabacum (common tobacco)
 C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 29-Sep-1999
 C:Accession: T03714
 R:Facchini, P.J.; Chappell, J.
 Proc. Natl. Acad. Sci. U.S.A. 89, 11088-11092, 1992
 A:Title: Gene family for an elicitor-induced sesquiterpene cyclase in tobacco.
 A:Reference number: Z15024; MUID:93066390; PMID:1438319
 A:Accession: T03714
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-550 <PAC>
 A:Cross-references: EMBL:L04680; NID:g170342; PIDN:AA19216.1; PID:g505588
 A:Experimental source: strain NK326
 C:Genetics:
 A:Introns: 37/3; 126/1; 251/2; 324/2; 370/3; 452/3
 C:Function:
 A:Description: mediates the conversion of the isoprenoid intermediate farnesyl diphosph
 C:Superfamily: vetispiradiene synthase 1

 Alignment Scores:
 Pred. No.: 1,07e-51 Length: 550
 Score: 809.50 Matches: 179
 Percent Similarity: 53.10% Conservative: 112
 Best Local Similarity: 32.66% Mismatches: 208
 Query Match: 27.84% Indels: 49
 DB: 2 Gaps: 10

 US-09-938-956-6 (1-1632) x T03714 (1-550)
 QY 1 ATGAGACGATCCGGAACACTCAACCTCTCTGTTGGATGTCAACTTCATCCATCGCTT 60
 Db 14 ValArgProValAlaAspPheSerProSerLeuTrpGlyAspGlnPheLeuSerPheSer 33
 QY 61 CTCAGTGACTATAGGAGGACAAACAGTGTAGTGTAGGCTTCTGAGCTGTCTACATTTGGTG 120
 Db 34 IleAspAsnGlnValAlaGluLysTyrIle-----Tyr 44
 QY 121 AAGATCGAATCGAGAAAGAAACGATCAATTCGACAA----- 159
 Db 45 AlaGlnGluIleGluAlaLeuLysGluGlnThrArgSerMetLeuLeuAlaThrGlyArg 64
 QY 160 -----CTTGAGTTGATCGATGCTTTCAGAGGATGGGCTGTCCGATCAT 204
 Db 65 LysLeuAlaAspThrLeuAsnLeuIleAspIleIleGluArgLeuGlyIleSerTyrHis 84
 QY 205 TTCGAAATGAGTTCGAAAGAAATCTGTCTCTATATATCTCGACCATCCTATTCACAG 264
 Db 85 PheGluLysGluIleAspGluIleLeuAspGlnIleTyrAsnGlnAsn----- 100
 QY 265 AACCTTTTCCAAAAGAAAGAGGATCTCTACTCCACATCTCTTCGATTTAGGCTCTC 324
 Db 101 -----SerAsnCysAsnAspLeuCysThrSerAlaLeuGlnPheArgLeuLeu 116
 QY 325 AGAGAACATGCTTTTCAAGTCCACAGAGGTATTTCGATAGTTTCAAGAACGAGGAGGT 384
 Db 117 ArgGlnHisGlyPheAsnIleSerProGluIlePheSerLysPheGlnAspGluAsnGly 136
 QY 385 GAGTTCAAAGAAAGCTTTAGCGACGACACGAGGATTTGCAACTGTATGATCAAGCTTCC 444
 Db 137 LysPheLysGluSerLeuAlaSerAspValLeuGlyLeuLeuAsnLeuTyrGluAlaSer 156
 QY 445 TTTCTGTTGACGGAAGGCGAAACACGCTCGAGTCAGGAGGGAATTCGCCACCAATTT 504
 Db 157 HisValArgThrHisAlaAspAspIleLeuGluAspAlaLeuAlaPheSerThrIleHis 176
 QY 505 TTGGAGGAAAAAGTGAACGAGGCTGTGTGATGCGGACCTTTTAAACAGATCGCATAT 564
 Db 177 LeuGluSerAlaAlaProHis-----LeuLysSerProLeuArgGluGlnValThrHis 194

565 TCTTTGGACATCCCTCTTCATTTGGAGGATTAAGGCGCAATGCACTGTGTGGATC--- 621
 195 AlaLeuGluGlnCysLeuHisGlyValProArgValGluThrArgPhePheLeuSer 214
 622 GAATGATATAGGAAGAGGCGGACATGATCCAGTAGTGTGGAGCTTGCATCTACGAC 681
 215 SerIleTyrAspGluGlnSerLysAsnValLeuLeuArgPheAlaLysLeuAsp 234
 682 TTAATATTGTTCAAGCACAATTTCAAGAAGAGCTCAAGAATCTTCAGGTGGGAGA 741
 235 PheAsnLeuLeuGlnMetLeuHisLysGlnLeuAlaGlnValSerArgTrpTrpLys 254
 742 AATACCTGGTGTCTGAGAAGTGGCTTCCGCAAGGATAGACTGCTGTAATCTACTTT 801
 255 AspLeuAspPheValThrThrLeuProTyrAlaArgAspArgValValGluCysTrpPhe 274
 802 TGGATPACTGGGATCATCGAGCCACGTCAGCATGCAAGTGCAGGATTAATGATGGCAAA 861
 275 TrpAlaLeuGlyValTyrPheGluProGlnTyrSerGlnAlaArgValMetLeuValLys 294
 862 GTCAACGCTGATTTACGCTGATCGATGATATTTATGATGTATGTCGACCTTAGAAGA 921
 295 ThrIleSerMetIleSerIleValAspPheThrPheAspAlaTyrGlyThrValLysGlu 314
 922 CTGGAACAATTCATCCTCAGCTTCAAGATGGATATAAATCAATCAATCGACCACTGCC 981
 315 LeuGluAlaTyrThrAspAlaIleGlnArgTrpAspIleAsnGluIleAspArgLeuPro 334
 982 GATTACATGCACTGCTTCTTGTGCTCACTCAACACTCTCGTGCATGATACATCGTACAT 1041
 335 AspTyrMetLysIleSerTyrLysAlaIleLeuAspLeuTyrLysAspTyrGluLysGlu 354
 1042 GTTATGAGGAGAAAGCGCTCAAGTATATACCTTACTCGCGCAATCGTGGTGTGATTG 1101
 355 LeuSerSerAlaGlyArgSerHisIleValCysHisAlaIleGluMetLysGluVal 374
 1102 GCGGATAGATATGATGAGGACGCGTCTTCTGCTCGGCGGCGGACCAACCAATTTGGA 1161
 375 ValArgAsnTyrAsnValGluSerThrTrpPheIleGluGlyTyrMetProValSer 394
 1162 GATGATTTGAGAACTCATGCGCAGTCGATGAAGTGGCCCTGTATGTTAAACGACATATTC 1221
 395 GluTyrLeuSerAsnAlaLeuAlaThr-----ThrThrTyrTyr 407
 1222 TTCGAGTAAACAGATTCGTTC-----ACAAAGAGACCGTCGACAGT 1263
 408 TyrLeuAlaThrThrSerTyrLeuGlyMetLysSerAlaThrGluGlnAspPheGluTrp 427
 1264 TTGTACAAATACCAAGATTTAGTTGTTGTCATCTCTGTTCTGGCGCTTCTGATGAT 1323
 428 LeuSerLysAsnProLysIleLeuGluAlaSerValIleCysArgValIleAspAsp 447
 1324 TTGGGAACCTCGGTGGAAGAGTGCAGCAGAGGGGATGTGCGCAATCACTTCAGTGCTAC 1383
 448 ThrAlaThrTyrGluValGluLysSerArgGlyGlnIleAlaThrGlyIleGluCysCys 467
 1384 ATGATGATCAATGATCGAGCGGCGGCGGCGGACGACGATGAATGGCTGATGCG 1443
 468 MetArgAspTyrGlyIleSerThrLysGluAlaMetAlaLysPheGlnAsnMetAlaGlu 487
 1444 GAGGTGGAAGATGATGCGGAGAGGCTGCGAAGGATTCCTCCATTCGCGCAAGAT 1503
 488 ThrAlaTrpLysAspIleAsn---GluGlyLeuLeuArgProThrProValSerThrGlu 506
 1504 TTTATAGGATGTCAGTTGATTTAGGAGGATGCGGAGTGTGATGATC---CATATGGA 1560
 507 PheLeuThrProIleLeuAsnLeuAlaArgIleValGluValThrTyrIleHisAsnLeu 526
 1561 GATGGGCGGCGCACACCAACCCCT 1584
 527 AspGlyTyr-----ThrHisPro 532

RESULT 8
 S68366
 (+)-delta-cadinene synthase isozyyme XC14 - Gossypium arboreum
 C:Species: Gossypium arboreum
 C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 29-Sep-1999
 C:Accession: S68366
 R:Chen, X.Y.; Chen, Y.; Heinstein, P.; Davissson, V.J.
 Arch. Biochem. Biophys. 324, 255-266, 1995
 A>Title: Cloning, expression, and characterization of (+)-delta-cadinene synthase: a c
 A:Reference number: S68365; MUID:96132653; PMID:8554317
 A:Accession: S68366
 A:Molecule type: mRNA
 A:Residues: 1-554 <CHE>
 A:Cross-references: EMBL:U3205; NID:G1045313; PIDN:AAA93065.1; PID:G1045314
 A:Experimental source: cultivar Nanking
 C:Superfamily: vetispiradiene synthase 1
 C:Keywords: phytoalexin biosynthesis

Alignment Scores:
 Pred. No.: 1,06e-50 Length: 554
 Score: 796.00 Matches: 185
 Percent Similarity: 52.07% Conservative: 104
 Best Local Similarity: 33.33% Mismatches: 232
 Query Match: 27.37% Indels: 34
 DB: 2 Gaps: 7

US-09-938-956-6 (1-1632) x S68366 (1-554)

QY 1 ATGAGACGATCCGAAACTACAAACCTTCTCGTTGG----- 36
 Db 21 MetArgProLysAlaAspPheGlnProSerIleTyrGlyAspLeuPheLeuAsnCysPro 40
 QY 37 GATGTCACATTCATCCATCGCTTCTCAGTACGATTAAGGAGGACAAACACGATGATAGG 96
 Db 41 AspLysAsnIleAspAlaGluThrGluLysArgHisGlnLeuLysGluValArg 60
 QY 97 GCTTCTGAGTGTGCTTGTGTAAGATGCACTGGAAGAAACGAGCAATCAATTCGA 156
 Db 61 -----LysMetIleValAlaProMetAlaAsnSerThrGln 72
 QY 157 CAACTGAGTGTGATGATGCTGAGAGATGGGCTGCGATCATTTCCAAATGAG 216
 Db 73 LysLeuAlaPheIleAspSerValGlnArgLysGlyValSerTyrHisPheThrLysGlu 92
 QY 217 TTCAAGAATCTGTCTCTATATATCTGACCATCACTATTACAGAAGCCCTTTTCCA 276
 Db 93 IleGluAspGluLeuGluAsnIleTyrHisAsnAsn----- 105
 QY 277 AAAGAAGAAAGGATCTCTACTCCACATCTCTTGCATTTAGGCTCTCAGAGAACATGGT 336
 Db 106 AspAlaGluAsnAspLeuTyrThrSerLeuArgPheArgLeuLeuArgGluHisGly 125
 QY 337 TTTCAAGTCCGACAAAGGATTCGATAGTTTCAAGAACGAGGAGGTGAGTTCAAGAA 396
 Db 126 PheAsnValSerCysAspValPheAsnLysPheLysAspGluGlnGlyAsnPheLysSer 145
 QY 397 AGCTTTAGCAGCAGACACGAGGATTTCTCAACTGTATGAAGCTTCTTCTTCTTTCAGC 456
 Db 146 SerValThrSerAspValArgGlyLeuLeuGluLeuTyrGlnAlaSerTyrLeuArgVal 165
 QY 457 GAAGCGCAACACCGCTGAGTCAGCGAGGGAATTCGCCACCAATTTTGGAGGAAAA 516
 Db 166 HisGlyGluAspIleLeuAspGluAlaIleSerPheThrThrAsnHisLeuSerLeuAla 185
 QY 517 GTGACGAGGCTGTGTGTGTGAGCGACCTTTTAAACAAGATTCGATATTTCTTTCGACATC 576
 Db 186 Val-----AlaSerLeuAspTyrProLeuSerGluValSerHisAlaLeuLysGln 203
 QY 577 CCTCTTATTCGGAGGATTAAGGCGCAATGCACCTGTGTGGATCGAATGATGATAGGAG 636
 Db 204 SerIleArgGlyLeuProArgValGluAlaArgHisTyrLeuSerValTyrGlnAsp 223
 QY 637 AGGCCGACATGAATCCAGTAGTGTGGAGCTTGCCATCTCGACTTCAATATTTGTTCAA 696

```

Db      224 IledgluSerHisAsnLysValLeuLeuGluPheAlaLysIleAspPheAsnMetValGln 243
QY      597 GCACAAATTCAGAGAGCTCAAGAAATCTTCAGGTGGTGGAGAAATAGCTGGTTTGT 756
Db      244 LeuLeuHisArgLysGluLeuSerGluIleSerArgTrpTrpLysAspLeuAspPheGln 263
QY      757 GAGAAGCTGCCCTTCGCAAGGATAGACTGGTGGAACTGCTACTTTGGATACTGGATC 816
Db      264 ArgLysLeuProTyrAlaArgAspArgValValGluGlyTyrPheTrpIleSerGlyVal 283
QY      817 ATCGAGCCAGCTCAGCATGCAAGTCAAGTAAATGATGGCGCAAGTCAACGCTCTGAT 876
Db      284 TyrPheGluProGlnTyrSerLeuGlyArgLysMetLeuThrLysValIleAlaMetAla 303
QY      877 ACGGTGATCATGATATTTATGATGCTGACCTTGAAGCACTTGAAGCACTTGAAGCACT 936
Db      304 SerIleValAspThrTyrAspSerTyrAlaThrTyrGluGluLeuIleProTyrThr 323
QY      937 GACCTCATCGAAGTGGATATAAATCAATCGACCAACTTCCGATTAACGCAACTG 996
Db      324 LysAlaIleGluArgTrpAspIleLysCysIleAspGluLeuProGluTyrMetLysPro 343
QY      997 TGTCTTCTTCGACTCAACATCTCGTCGATGATACATCGATGATGATGATGATGATG 1056
Db      344 SerTyrLysAlaLeuLeuAspValTyrGluGluMetGluGlnLeuValAlaLysHisGly 363
QY      1057 GCGTCAACGTTATACCTTACCTGCGCAATCGTGGTGTGATTTGGCGGATAAGTATATG 1116
Db      364 ArgGlnTyrArgValGluTyrAlaLysAsnAlaMetIleArgLeuAlaGlnSerTyrLeu 383
QY      1117 GTAGAGGACGCTGCTTACGGGGGCAAAACCAAGTTTGGAGAGTATTGGAGAAC 1176
Db      384 ValGluAlaArgTrpThrLeuGlnAsnTyrLysProSerPheGluGluPheLysAlaAsn 403
QY      1177 TCATGCGACGTCAATAAGTGGCGCTGTATGTTAAACGCACATATCTTCGAGTAACAT 1236
Db      404 AlaLeuProThrCysGlyTyrAlaMetLeuAlaIleThrSerPheValGlyMetGlyAsp 423
QY      1237 TCGTTCAAAAGAGACCGCTGCACAGTTGTGTACAAATACCAGTATTAGTTGCTGTGCA 1296
Db      424 IleValThrProGluThrPheLysTrpAlaAlaAsnAspProLysIleIleGlnAlaSer 443
QY      1297 TCCTTCGTTCTCGCTTCCTGATCATTTGGAACTCGGTGGAGAGAGTGGAGAGAGG 1356
Db      444 ThrIleIleCysArgPheMetAspValAlaGluHisLysPheLysHisArgArgGlu 463
QY      1357 GATGTGCCGAATCACTTCAGTGTACATGACTGACTCAATGCAATGCGAGCGGCGAGCG 1416
Db      464 AspAspCysSerAlaIleGluCysTyrMetGluGluTyrGlyValThrAlaGlnGluAla 483
QY      1417 -----CGAAGCAGCTGAAATGGCTGATAGCGGAGGTGTGGAAGAAGATGAAT 1464
Db      484 TyrAspValPheAsnLysHisValGlu-----SerAlaTrpLysAspValAsn 499
QY      1465 GCGAGAGGGTGTCCGAAGATTCTCCATTCCGCAAGATTTTATAGGATGCGAGTTGAT 1524
Db      500 LysGluPheLeu---LysProThrGluMetProThrGluValLeuAsnArgSerLeuAsn 518
QY      1525 TTAGAAGAGTGGCGCATGTTGATGACATATGAGATGGCGGAGCGGCACACACACCT 1584
Db      519 LeuAlaArgMetAspValLeuTyrArgGluGlyAspGlyTyrThrTyrValGlyLys 538
QY      1585 ATTATACATCAACAAATCACCAGAACCTTATTCGAGCCCTTTCGA 1629
Db      539 AlaAlaLysGlyIleThrSerLeuLeuIleGluProValAla 553

```

RESULT 9

S68365

[+]-delta-cadinene synthase isozyme XC1 - Gossypium arboreum

C:Species: Gossypium arboreum

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 29-Sep-1999

C:Accession: S68365

R:Chen, X.Y.; Chen, Y.; Heinsteins, P.; Davison, V.J.
 Arch. Biochem. Biophys. 324, 255-266, 1995
 A:Title: Cloning, expression, and characterization of (+)-delta-cadinene synthase: a c
 A:Reference number: S68365; MUID:96132653; PMID:8554317
 A:Accession: S68365
 A:Molecule type: mRNA
 A:Residues: 1-554 <CH>
 A:Cross-references: EMBL:U22206; NID:G1045311; PID:AAA93064.1; PID:G1045312
 A:Experimental source: cultivar Nanking
 C:Superfamily: vetispiradiene synthase 1
 C:Keywords: phytoalexin biosynthesis

Alignment Scores:
 Pred. No.: 3,158-49 Length: 554
 Score: 776.00 Matches: 178
 Percent Similarity: 52.07% Conservative: 111
 Best Local Similarity: 32.07% Mismatches: 232
 Query Match: 26.69% Indels: 34
 DB: 2 Gaps: 7

US-09-938-956-6 (1-1632) x S68365 (1-554)

```

QY      1 ATGAGACGATCCGAAACTACACCTTCTCGTTGG----- 36
Db      21 MetArgProLysAlaAspPheGlnProSerIleTrpGlyAspLeuPheLeuAsnCysPro 40
QY      37 GATGTCAACTTCACCAATCGTCTCAGTCACTATAGGAGGACAAACACGATGATGAG 96
Db      41 AspLysAsnIleAspAlaGluThrGluLysArgHisGlnGlnLeuLysGluValArg 60
QY      97 GCTTCTGAGTGGTCACTTGTGGTGAAGATCGAATCGAAGAGAAAGAAACGATCAAAATCGA 156
Db      61 -----LysMetIleValAlaProMetAlaAsnSerThrGln 72
QY      157 CACTTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 216
Db      73 LysLeuAlaPheIleAspSerValGlnArgLeuGlyValSerTyrHisPheThrLysGlu 92
QY      217 TCAAGAAATCTTGTCTCTATATATCTGACCATCATTTACAGAACCCCTTTTCCA 276
Db      93 ILeGluAspGluLeuGluAsnIleTyrHisAsnAsn----- 105
QY      277 AAGAAGAAGAGGATCTTACTCCATCTCTTGCATTTTAGGCTCCTCAGAGAACATGTT 336
Db      106 AspAlaGluAsnAspLeuTyrThrSerIleArgPheArgLeuLeuArgGluHisGly 125
QY      337 TTTCAAGTCCCAAGAGCTATTCGATAGTATTTCAAGACAGAGAGGAGGTGATTCAGAA 396
Db      126 TyrAsnValSerCysAspValPheAsnLysPheLysAspGluGlnGlyAsnPheLysSer 145
QY      397 AGCTTAGCGACGACACACAGAGATGTTGCAACTGTATGAAGCTTCTTCTTGTGACG 456
Db      146 SerValThrSerAspValArgGlyLeuLeuGluLeuTyrGlnAlaSerTyrLeuArgVal 165
QY      457 GAAGGCGAACAACCGTCCAGTCAGGAGGGAATTCGCCCAAAATTTTGGAGGAAAA 516
Db      166 HisGlyGluAspIleLeuAspGluAlaIleSerPheThrHisIleLeuSerLeuAla 185
QY      517 GTGAACGAGGTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 576
Db      186 Val-----AlaSerLeuAspHisProLeuSerGluGluValSerHisAlaLeuLysGln 203
QY      577 CCTCTTCATGAGGATTAAGAGCAATGACCTGTGTGATGATGATGATGATGATGATG 636
Db      204 SerIleArgArgGlyLeuProArgValGluAlaArgHisTyrLeuSerValTyrGlnAsp 223
QY      637 AGGCCCGACATCAATCCAGTAGTGTGGAGCTTGCATCTGATGATGATGATGATGATG 696
Db      224 IleGluSerHisAsnLysAlaLeuLeuGluPheAlaLysIleAspPheAsnMetLeuGln 243
QY      697 GCACAAATTCAGAGAGCTCAAGAAATCTTTCAGGTGGTGGAGAAATACTGGTGTGTT 756
Db      244 PheLeuHisArgLysGluLeuSerGluIleCysArgTrpTrpLysAspLeuAspPheGln 263

```

QY 757 GAGAGCTGCCCTTCGCAAGGATAGACTGGTGGAAATGCTACTTTTGGAAATACCTGGGATC 816
 Db 264 ARGLYSLeuProTyrAlaArgAspArgValValGluGlyTyrPheTrpIleSerGlyVal 283
 QY 817 ATCGAGCCAGTCAGATGCAAGTGCAGGATATGATGGGCAAGTCAACGCTCTGATT 876
 Db 284 TyrPheGluProGlnTyrSerLeuGlyArgLysMetLeuThrLysValIleAlaMetAla 303
 QY 877 ACGGTGATCGATGATATTATCATGCTATGACCTTAGAAGAACTCGAACAAATTCAT 936
 Db 304 SerIleValAspThrTyrAspSerTyrAlaThrTyrGluGluIleProTyrThr 323
 QY 937 GACCTCATTCGAGATGGATATAAACAATCGACCAACTTCCGATATCATGCAACTG 996
 Db 324 AsnAlaIleGluArgTrpAspIleLysCysIleAspGluIleProGluTyrMetLysPro 343
 QY 997 TCGTTCTTCGCACTCAACACTTCGTCGATGATACATCGTACGATGTTTATGAAGAGAAA 1056
 Db 344 SerTyrLysAlaLeuLeuAspValTyrGluGluMetValGlnLeuValAlaGluHisGly 363
 QY 1057 GCGCTCAACGTTATACCTTACCTCGCGCAATCGTGGTGTGATTTGGCGGATAGTATATG 1116
 Db 364 ArgGlnTyrArgValGluTyrAlaLysAsnAlaMetIleArgLeuAlaGlnSerTyrLeu 383
 QY 1117 GTAGAGGACGCTGCTTACCGCGGCGACAAACCAAGTTTGGAGAGTATTTGGAGAAC 1176
 Db 384 ValGluAlaLysTrpThrLeuGlnAsnTyrLysProSerPheGluGluPheLysAlaAsn 403
 QY 1177 TCATGCGACGTCGATAGTGGCGCTGTATGTTAAAGCACAATATCTCCGAGTAAACAGAT 1236
 Db 404 AlaLeuProThrCysGlyTyrAlaMetLeuAlaIleThrSerPheValGlyMetGlyAsp 423
 QY 1237 TCGTTCAAAAGAGACCGTCGACAGTTTGTACAATACCAAGATTTAGTTCGTGTCA 1296
 Db 424 IleValThrProGluThrPheLysTrpAlaAlaSerAspProLysIleIleGlnAlaSer 443
 QY 1297 TCCTCGTCTCGCGCTTCTGATGATTTGGCAACCTCGTGGAGAGGTGACGAGAGG 1356
 Db 444 ThrIleLysCysArgPheMetAspValAlaGluHisLysPheLysHisArgGlu 463
 QY 1357 GATGTCGCGAAATCATCTCAGTGCTCATGATGATGATGATGATGATGATGATGATGATG 1416
 Db 464 AspAspCysSerAlaIleGluCysTyrMetGluGluTyrGlyValThrAlaGlnGluAla 483
 QY 1417 -----CGGAGCAGCTGAATGCTGATAGCGAGGTGGAGAGATGAAT 1464
 Db 484 TyrAspValPheAsnLysHisValGlu-----SerAlaTrpLysAspLeuAsn 499
 QY 1465 GCGGAGAGGTGTCGAAGGATTTCCATTCGGCAAGATTTTATAGATGTCAGTTGAT 1524
 Db 500 GlnGluPheLeu--LysProThrGluMetProThrGluValLeuAsnArgSerLeuAsn 518
 QY 1525 TTAGGAGAGTGGCGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1584
 Db 519 LeuAlaArgValMetAspValLeuTyrArgGluGlyAspGlyTyrThrTyrValGlyLys 538
 QY 1585 ATTATACATCAACAAATGACCAAGCTTATTCGAGCCCTTTGCA 1629
 Db 539 AlaAlaLysGlyGlyIleThrSerLeuLeuIleGluProIleAla 553

RESULT 10
 T06266
 germacrene C synthase - tomato
 C/Species: Lycopersicon esculentum (tomato)
 C/Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 11-May-2000
 R/Accession: T06266
 R/Colby, S.M.; Crook, J.; Dowdle-Rizzo, B.; Lemaux, P.G.; Croteau, R.
 Proc. Natl. Acad. Sci. U.S.A. 95, 2216-2221, 1998
 A/Title: Germacrene C synthase from Lycopersicon esculentum cv. VFNT Cherry tomato: cDNA
 A/Reference number: Z15576; MUID:98151492; PMID:9482865
 A/Accession: T06266
 A/Status: preliminary; translated from GE/EMBL/DBJ

A;Molecule type: mRNA
 A;Residues: 1-548 <COL>
 A;Cross-references: EMBL:AF035631; NID:g2967688; PIDN:AAC39432.1; PID:g2967689
 A;Experimental source: cultivar VFNT
 C;Superfamily: vetispiradiene synthase 1

Alignment Scores:
 Pred. No.: 3,92e-46 Length: 548
 Score: 734.00 Matches: 177
 Percent Similarity: 50.72% Conservative: 106
 Best Local Similarity: 31.72% Mismatches: 233
 Query Match: 25.24% Indels: 42
 Db: 2 Gaps: 11

US-09-938-956-6 (1-1632) x T06266 (1-548)

QY 4 AGAGCATCCGAAACATACCAACCTTCCTCGTGGGATGTCACATTCATCCATCGCTT--- 60
 Db 10 ArgProLeuAlaAsnPheHisProSerValTyrGlyTyrHisPheLeuSerTyrThrHis 29
 QY 61 -----CTCAGTGACTATAAGGAGGACAAACACGCTGATT 93
 Db 30 GluIleThrAsnGlnGluLysValGluValAspGluTyrLysGlu----- 44
 QY 94 AGGCTTCCTGAGCTGCTCATTCTTGTGAAGATGAACTCGAGAAAGAAACGATCAAAAT 153
 Db 45 -----ThrIleArgLysMetLeuValGluThrCysAspAsnSerThr 58
 QY 154 CGACACTTCAGTGTGATCGATGATGTCAGAGATGGGCTGTCGATCATTTCCAAAT 213
 Db 59 GlnLysLeuValLeuIleAspAlaMetGlnArgLeuGlyValAlaTyrHisPheAspAsn 78
 QY 214 GAGTTCAAAGAAATCTTGTCTCTATATATCTGACCATCATCTATTACAGAACCCCTTTT 273
 Db 79 GluIleGluThrSerIleGlnAsnIlePhe--AspAlaSerSerLysGlnAsn----- 95
 QY 274 CCAAGAGAAAGGATCTCTACTCCACATCTCTGATTTAGGCTCTCTCAGAGACAT 333
 Db 96 ---AspAsnAspAsnLeuTyrValSerLeuArgPheArgLeuValArgGln 114
 QY 334 GGTTCCTCACTCGCACAGAGGTATTCATAGTTTCAAGAACGAGGAGGTGATTCAAA 393
 Db 115 GlyHisTyrMetSerSerAspValPheLysGlnPheThrAsnGlnAspGlyLysPheLys 134
 QY 394 GAAACCTTCAGCAGCAGCAGAGGATTTGCACTGTATGAGTCTCTCTCTCTGTTG 453
 Db 135 GluThrLeuThrAsnAspValGlnGlyLeuLeuSerLeuTyrGluAlaSerHisLeuArg 154
 QY 454 ACGGAGCGCAACCCAGCTCGAGTCAGCGAGGAAATTCGCCACCAAAATTTTGGAGAA 513
 Db 155 ValArgAsnGluGluIleLeuGluGluAlaLeuThrPheThrThrHisLeuGluSer 174
 QY 514 AAGTG---AACGAGGTGTTGATGGGACCTTTTAAACAAGATCGCATATCTTTG 570
 Db 175 IleValSerAsnLeuSerAsnAsnAsnSerLeuLysValGluValGlyGluAlaLeu 194
 QY 571 GACATCCCTCTTCATTGAGGATTTAAAGGCCAAATTCACCTGTGTGGATCGAATGGTAT 630
 Db 195 ThrGlnProIleArgMetThrLeuProArgMetGlyAlaArgLysTyrIleSerIleTyr 214
 QY 631 AGGAGAGGCGCCGACATGAATCCAGTAGTGTGAGCTTGGCATACTCGACTTAAATATT 690
 Db 215 GluAsnAsnAspAlaHisHisLeuLeuLysPheAlaLysLeuAspPheAsnMet 234
 QY 691 GTTCAAGCACAAATTTCAAGAGAGCTCAAGAGATTCCTTCAGGTGGTGGAGAAATCTGGG 750
 Db 235 LeuGlnLysPheHisGlnArgGluLeuSerAspLeuThrArgTrpTyrLysAspLeuAsp 254
 QY 751 TTTTGTGAGAAAGTCCCTTCGCAAGGATAGACTGTGGAAATCTCTCTTTGGAATCT 810
 Db 255 PheAlaAsnLysTyrProTyrAlaArgAspArgLeuValGluCysTyrPheTrpIleLeu 274
 QY 811 GGGATCATCGAGCCACGTCAGCATCGACATGCAAGGATTAATGATGGGCAAGTCAACGCT 870


```

QY 973 CAACCTCCGATTACATCAACTGCTGCTTCTTCTGCACTCAACAACCTGCGTATGATACA 1032
Db 258 ThrLeuProLysTyrMetArgValCysPheGluAlaLeuAspMetIleThrThrGluIle 277
QY 1033 TCGTACGATGTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1089
Db 278 SerMetLysIleTyrLysSerHisGlyTrpAsnProThrAlaLeuArgGlnSerVal 297
QY 1090 -----TGGGTGATTGGCGGATAGATATATGTTAGAGAGAGAGAGAGAGAGAGAG 1134
Db 298 IleIleIleGlnAsnTrpAlaSerLeuCysLysAlaPheLeuValGluAlaLysTrpPhe 317
QY 1135 TACGCGCGGACAAACCAAGTTTGAAGAGATTTGGAGAACTCATGGCAGTCCGTAAGT 1194
Db 318 AsnSerGlyTyrLeuProThrThrGluGluTyrMetLysAsn----- 331
QY 1195 GGGCCCTGTATGTTAAACGCACATATTTCTCCGAGTAAACAGATTCGTTCCAAAGGAGACC 1254
Db 331 ----- 331
QY 1255 GTCGACAGTTGTCACAAATACACAGATTAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1314
Db 331 ----- 331
QY 1315 GCTGATGATTGCGGAACCTCGTGAAGAGGTGACGACAGAGGAGTGTCCGAAATCACTT 1374
Db 332 -----GlyValAspGluAsnGlnAspGlyThrAspGlySerTyrVal 345
QY 1375 CAGTCTACATGAGTACTACATCAATGCATCGAGGCG-----GAGGCGCGGAGACAGTGA 1431
Db 346 GluCysTyrLeuAsnGluTyrLysGlySerThrValAspGluAlaArgThrHisValala 365
QY 1432 TGGCTGATAGCGGAGGTGGAGAGAGATGAATGCGGAGAGGTGTCGAAGGATTCCTCA 1491
Db 366 GlnLysIleSerArgAlaTrpLysArgLeuAsnArgGluCysLeuAsn---ProCysPro 384
QY 1492 TTGCGCAAGATTTTATAGATGTCGATGTTGATTTAGGAAGAGTGGCGCATGTTGATGAC 1551
Db 385 PheSerArgSerPheSerLysAlaCysLeuAsnIleAlaArgThrValProLeuMetTyr 404
QY 1552 CATATGAGGAT 1563
Db 405 SerTyrAspAsp 408

RESULT 15
C56118
vetispiradiene synthase 2 - Hyoscyamus muticus (fragment)
C:Species: Hyoscyamus muticus
C:Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 29-Sep-1999
C:Accession: C56118
R:Back, K.; Chappell, J.
J. Biol. Chem. 270, 7375-7381, 1995
A:Title: Cloning and bacterial expression of a sesquiterpene cyclase from Hyoscyamus muticus
A:Reference number: A56118; MUID:95221394; PMID:7706281
A:Accession: C56118
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-350 <BA>
A:Cross-references: GB:U20190; NID:G763426; PIDN:AA86339.1; PID:G763427
A>Note: authors translated the codon GAA for residue 181 as Val, and GCA for residue 182
C:Superfamily: vetispiradiene synthase 1

Alignment Scores:
Pred. No.: 1-55e-33 Length: 350
Score: 562.50 Matches: 116
Percent Similarity: 53.64% Conservative: 68
Best Local Similarity: 33.82% Mismatches: 140
Query Match: 19.34% Indels: 19
DB: 2 Gaps: 5

US-09-938-956-6 (1-1632) x C56118 (1-350)
QY 580 CITTCATTGGAGGATTAAAGGCCAAATGCACCTGTGTGATCGAATGGTATAGGAAGAGG 639

```

Search completed: February 25, 2004, 05:50:38

```

Db 1 LeuHisLysSerIleProArgValGluThrArgTyrPheIleThrIleTyrGluGlu 20
QY 640 CCGACATCAATCCAGTACTGTTGGAGCTTGGCATACTCGACTTAATAATTGTTCAAGCA 699
Db 21 GluLeuLysAsnAspValLeuLeuArgPheAlaLysLeuAspPheAsnLeuLeuGlnMet 40
QY 700 CAATTTCAAGAGAGCTCAAGAACTCTCAGTGGTGGAGAAATACTGGGTTGTTGAG 759
Db 41 LeuHisLysGlnGluLeuThrGluValSerMetTrpTrpLysAspLeuAspPheValThr 60
QY 760 AAGCTCCCTTCGCAAGGATAGACTGGTGGATGCTACTTTTGGTAATACTGGGATCATC 819
Db 61 ThrLeuProTyrAlaArgAspArgAlaValGluCysTyrPheTrpThrValGlyValTyr 80
QY 820 GAGCCAGTCCAGATGCAAGTCCAGGATATGATGGGCAAGATCAACCTCTGATTACG 879
Db 81 AlaGluProGlnTyrSerGluAlaArgValMetLeuAlaLysThrIleAlaMetIleSer 100
QY 880 GTCATCGATGATATTATGATGCTATGCACCTCTAGAAAGAACTCGAAACAATTCACTGAC 939
Db 101 IleValAspAspThrPheAspAlaTyrGlyIleValLysGluLeuGluValTyrThrAsp 120
QY 940 CTCATTCAAGATGGGATATAAATCAATCGACCAACTCCGATTCATGCAACTGTGC 999
Db 121 AlaIleGlnArgTrpAspIleAsnGlnIleAspArgLeuProAspTyrMetLysIleSer 140
QY 1000 TTTCTTGCCTCAACAACCTTCGTCGATGATACATCCTAGATGTTATGAGGAGAAAGGC 1059
Db 141 TyrLysValLeuLeuAspLeuTyrLysAspTyrGluThrGluLeuSerLysAspGlyArg 160
QY 1060 GTCACAGTTATACCTACCTCGCGCAATCGTGGTGTGATTTGGCGGATAGTATATGTA 1119
Db 161 SerGluValValHisTyrAlaLysGluArgMetLysGluIleValArgAsnTyrPheVal 180
QY 1120 GAGGACGGTGGTCTACGCGGCGCAACAAGTTTGAAGAGATTTGGAGAACTCA 1179
Db 181 GluAlaLysTrpPheIleGluGlyTyrMetProProValSerGluTyrLeuAsnAsnArg 200
QY 1180 TGGCAGTCCGATAAGTGGCGCTGTATGTTAAACGACATATTTCCGAGTAACAGATTCG 1239
Db 201 -----LeuAlaThrSerThrTyrTyrLeuLeuThrThrThr 212
QY 1240 -----TTCACAAAGGAGACCGTCGACAGTTTGTACAAATACCAC 1278
Db 213 SerTyrLeuGlyMetLysCysAlaAsnLysGluAspPheGluTrpLeuThrLysAsnPro 232
QY 1279 GATTAGTTCTGTTGTCATCTTCGTTCTCGCGCTTCTGCTGATGATTTGGAACTCGGTG 1338
Db 233 LysIleLeuGluAlaAsnValThrLeuCysArgValIleAspAspIleAlaThrTyrGlu 252
QY 1339 GAGAGGTGAGCAGAGGGATGTCGGAATCACTTCAGTGTCTACATGACTGACTACAT 1398
Db 253 ValGluLysGlyArgGlyGlnIleAlaThrGlyIleGluCysTyrMetArgAspTyrGly 272
QY 1399 GCATCGGAGCGGCGGCGGAGACCGTCGAAATGCTGATAGCGAGGTGTGGAAGAAG 1458
Db 273 ValSerThrGluGluAlaMetGluLysPheGlnGluMetAlaGluIleAlaTrpLysAsp 292
QY 1459 ATGAATCGGAGAGGGTGTGCAAGGATTTCCATTTCGCGCAAGATTTTATAGGATGCA 1518
Db 293 ValAsn---GluGlyIleLeuArgProThrProValSerAlaLysIleLeuThrArgIle 311
QY 1519 GTTCATTTAGGAAGGATGGCGCGAGTTCATGTCAC---CATAATGGAGATGGCGCGGACA 1575
Db 312 LeuAsnLeuAlaArgIleIleAspValThrTyrLeuHisAsnGlnAspGlyTyr----- 329
QY 1576 CAACACCT 1584
Db 330 ThrHisPro 332

```

Job time : 74 secs
